

76910

From: Rao, Manjunath N.
Sent: Wednesday, October 02, 2002 9:16 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/656,640

From: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10C 01
Phone: 306-5681

Dat : 10-2-02

Please search the following as soon as possible for application with serial number **09/656,640**

SEQ ID NO:1, and 3 against all commercial nucleic acid databases including issued patents database and pending application database and provide a **print of all results**.

SEQ ID NO:2 and 4, against all commercial protein databases including issued patents database and pending application database and provide a **print of all results**.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.
Bi technology Patent Examiner
Art Unit 1652, Room 10A11
Crystal Mall 1, USPTO.

RECEIVED
OCT - 2 2002
CH/DEPT. DIVISION
(STIC)

POINT OF CONTACT:
PAUL SCHULWITZ
TECHNICAL INFO. SPECIALIST
CM1 6B06 TEL. (703) 305-1954

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 10/3
Date Completed: 10/7
Searcher Prep/Review: 10
Clerical: _____
Online time: 10

TYPE OF SEARCH:
NA Sequences: 2
AA Sequences: 2
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

✓

THIS PAGE BLANK (USPTO)

RECEIVED
JAN 10 1964
U.S. PATENT OFFICE
WASHINGTON, D.C.

Matches 1958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ggtatcatacaatgatgaagcgaagctatcagagccgttgagccttgagccttgatc 60
DB 1 ggtatcatacaatgatgaagcgaagctatcagagccgttgagccttgatc 60
OY 61 ggcggcagctctgtcgtatgcagatccgtctgtctgtcgtatgcagatccgtct 120
DB 61 ggcggcagctctgtcgtatgcagatccgtctgtctgtcgtatgcagatccgtct 120
OY 121 ctcacacaaagagcagagcagcagctgagctcccttgagccttgagccttgagc 180
DB 121 ctcacacaaagagcagagcagcagctgagctcccttgagccttgagccttgagc 180
OY 181 atccctccctctgaagcgcgcgaagtaagtaacattctatagcttagagagcagct 240
DB 181 atccctccctctgaagcgcgcgaagtaagtaacattctatagcttagagagcagct 240
OY 241 tgcataatcgtcagtaacgttcccaaccccaacacacacacacacacacacacacac 300
DB 241 tgcataatcgtcagtaacgttcccaaccccaacacacacacacacacacacacacac 300
OY 301 gattgagatgaagcctctcccaacacacacacacacacacacacacacacacacac 360
DB 301 gattgagatgaagcctctcccaacacacacacacacacacacacacacacacacac 360
OY 361 tggatacagaatgagcctctcccaacacacacacacacacacacacacacacacacac 420
DB 361 tggatacagaatgagcctctcccaacacacacacacacacacacacacacacacacac 420
OY 421 tgcctcgtctgtgaacagcagagagacacacacacacacacacacacacacacacac 480
DB 421 tgcctcgtctgtgaacagcagagagacacacacacacacacacacacacacacacac 480
OY 481 tttctcgtcgtctgtcgtctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 540
DB 481 tttctcgtcgtctgtcgtctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 540
OY 541 gattactactacacacacacacacacacacacacacacacacacacacacacacac 600
DB 541 gattactactacacacacacacacacacacacacacacacacacacacacacacac 600
OY 601 gtcacacacacacacacacacacacacacacacacacacacacacacacacacac 660
DB 601 gtcacacacacacacacacacacacacacacacacacacacacacacacacacac 660
OY 661 ggcctgagatgagcctcgaagcctcgaagcctcgaagcctcgaagcctcgaagcct 720
DB 661 ggcctgagatgagcctcgaagcctcgaagcctcgaagcctcgaagcctcgaagcct 720
OY 721 gactgacacacacacacacacacacacacacacacacacacacacacacacacac 780
DB 721 gactgacacacacacacacacacacacacacacacacacacacacacacacacac 780
OY 781 cttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 840
DB 781 cttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 840
OY 841 aagatcagatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 900
DB 841 aagatcagatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 900
OY 901 tggcctatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
DB 901 tggcctatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
OY 961 tgaagcctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1020
DB 961 tgaagcctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1020
OY 1021 cagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
DB 1021 cagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
```

```
OY 1081 atctctatgcccagcgcctggagagtggtatcagacttcccaacttcgtcgcagctc 1140
DB 1081 atctctatgcccagcgcctggagagtggtatcagacttcccaacttcgtcgcagctc 1140
OY 1141 atcagatccgaacacacacacacacacacacacacacacacacacacacacacac 1200
DB 1141 atcagatccgaacacacacacacacacacacacacacacacacacacacacacac 1200
OY 1201 actgacaaagtgatgagatgctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1260
DB 1201 actgacaaagtgatgagatgctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1260
OY 1261 gtcgctcgaacacacacacacacacacacacacacacacacacacacacacacac 1320
DB 1261 gtcgctcgaacacacacacacacacacacacacacacacacacacacacacacac 1320
OY 1321 cccactgatacagagacacacacacacacacacacacacacacacacacacacacac 1380
DB 1321 cccactgatacagagacacacacacacacacacacacacacacacacacacacacac 1380
OY 1381 gttacctctcgtgagtgagagacacacacacacacacacacacacacacacacacac 1440
DB 1381 gttacctctcgtgagtgagagacacacacacacacacacacacacacacacacacac 1440
OY 1441 atctgagcagctgagagacacacacacacacacacacacacacacacacacacacacac 1500
DB 1441 atctgagcagctgagagacacacacacacacacacacacacacacacacacacacacac 1500
OY 1501 gacttcgagctcctctcgtctcgtctcgtctcgtctcgtctcgtctcgtctcgtctc 1560
DB 1501 gacttcgagctcctctcgtctcgtctcgtctcgtctcgtctcgtctcgtctcgtctc 1560
OY 1561 ctcgaagagtgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1620
DB 1561 ctcgaagagtgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1620
OY 1621 ccttcgcgttaagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1680
DB 1621 ccttcgcgttaagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1680
OY 1681 gttggtctcaatgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1740
DB 1681 gttggtctcaatgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1740
OY 1741 cttcaatgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1800
DB 1741 cttcaatgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1800
OY 1801 tggagcctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1860
DB 1801 tggagcctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1860
OY 1861 acttcagcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1920
DB 1861 acttcagcagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1920
OY 1921 cccagcctgatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1980
DB 1921 cccagcctgatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1980
```

RESULT 2
AX101096
LOCUS AX101096 1958 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 1 from Patent WO0121809.
ACCESSION AX101096
VERSION AX101096.1 GI:13619951
KEYWORDS
SOURCE Stachybotrys chartarum.
ORGANISM Stachybotrys chartarum.
REFERENCE 1 (bases 1 to 1958) Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Stachybotrys.

AUTHORS Wang, H.
TITLE Stachybotrys phenol oxidizing enzyme
JOURNAL Patent: WO 0121809-A 1 29-MAR-2001
GENCOR INTERNATIONA- INC- (us)
FEATURES
Source Location/Qualifiers
1. 1958
/organism="Stachybotrys chartarum"
/db_xref="taxon:74722"

BASE COUNT 394 a 593 c 490 g 481 t
ORIGIN

Query Match 100.0%; Score 1958; DB 6; Length 1958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggtacccatcaatgatcagcccaagctatcgagccgtgctcgtgacctgtgatac 60
DB 1 GGATCCATCAACGATGATCAGCCAAAGCTATCGACCGCTGCTGGGCTTGATC 60
QY 61 ggcggagcctctgcagatgcagatcctgtgtgtgctgctgagacagagcccttc 120
DB 61 GCGCGAGCTCTGTGATGCGAGATCCGTGTGTGTGATGAGACAGATCCCTTCGG 120
QY 121 ctcccaagagcagacgactgagctccctccctgtgctgtacgaagtgtcctgcg 180
DB 121 CTCACCAAGAGGAGCAGACGCTGATCTCCCTGGCCTTGTCGAAGTGCTTGCGG 180
QY 181 atccctctctgaaagggcccaagtaagtaattatctatagctagcagaagcaagt 240
DB 181 ATCCCTCTCTGAAGGGCCCAATGATGATGATCTATGAGCTGACAGACCAACGT 240
QY 241 tgcatacattgcagatcagctccccaacccaacacatgagagagacatctgtacta 300
DB 241 TGCATACATTCGATGACGTCCCTCCCAACCCCAACACTGAGAGACATCTTGTA 300
QY 301 gatgagatagagccctctcccaacacatctacacatctgagacggcccaacatgt 360
DB 301 GATGAGATTAGGCCCTTCTCCACAGATCTACCTGATGTGAGCGCGGCAACATGGT 360
QY 361 tggataagatgagctgtcccaagactacatcatctgttccctgtgtgacatgagat 420
DB 361 TGGATAGATGAGCTGTCCCAAGACTACCATCATGCTTCTGTGGCACTGAGATGT 420
QY 421 tgcagctctgtgagacgagagaaacacctctcccaacagctccacttgacagctc 480
DB 421 TGTCCGCTTGTGTAACGCGAGAGAACACTCTCCCAACAGCTCCACTTGACAGGCTC 480
QY 481 ttctctcagactccctttagtgtgtgagtgagacactaccagccttgagactaa 540
DB 481 TTTCTCTCGAGCTCCCTTGTGATGTTGGCTGAGGACATACCAAGCCTGGCGATCAA 540
QY 541 ggtattactactaccccaagcagcgtgcgcgactgttgttacaatgacatgcat 600
DB 541 GGATATTACTATACCCCAACAGGAGGCTGCCGATCTTTGGATCATATACATATGCAT 600
QY 601 gtccatcacgcgcagagacactacatgagtgagctgagtgatctacatgagagccc 660
DB 601 GTCCATACCGCGCAGAGAACGCTATCATGTGTCAGCTGCTGTACTGATCCAGAGACC 660
QY 661 ggtcgaagatgacctcaccagcagctacagcgaatgttatacccttggcttct 720
DB 661 GGTGAGAGATGACCTTGAACCTCCCAAGCGCTACGGGAGTTGATATCCCTTGTTCT 720
QY 721 gactgcgaagcgatatacagcagagcactctcttctcacaatgagaggtttcag 780
DB 721 GACTGCCAAGCGATACACGACGAGCTCTCTTCTTCCACCAATGAGAGGTTTCCAG 780
QY 781 ctctcgggtgagcgttataatgtgtaagtgtgagccattgagatcttcacatctag 840
DB 781 CTCTCGGGTGAGCTTATTTCAAGTGTAAAGTGAAGCCCATTTGAGATCTTCAGATCTTAG 840
QY 841 aagtatcgatgatatgaattgtgcatgctctaaccagtgctatcacagaaagctcagc 900

DB 841 AAGTATGATGTATGAATATGTGATGCTCTTAACCACTGCTATACAGAGAGGTCAACCT 900
QY 901 tggactatgtcacaagcgtgcagcgcgaagtagccgtctcgtctcccaagctgcgtc 960
DB 901 TGGCTATGTCTCAACGTGACGCGCGGAGATACGCTTCCGCTTCCCAACGCTGCGCTC 960
QY 961 tcaagcctctcgtctgtatctgtacactctgagagatccagagacagactcccttc 1020
DB 961 TCACGCTCTTTCGCTCTGTATCTTGTACTCTGAGGATTCAGAGACAGACTTCCCTTC 1020
QY 1021 caagltcatgcgcgtgacagctgctgtctgtgagggccctgtgacactgacatctgac 1080
DB 1021 CAGGTATGCGCGCTGACGAGGCTGCTGCTGAGGCGCCCTGTGACATGACACTCTGTAC 1080
QY 1081 atctctatgacgcagcgtgagaggtgtatgactgaactctccactctcgtgacagtc 1140
DB 1081 ATCTCTATGGCGGACCGCTGAGAGGTTGTATGCACTTCTCCACTTCCCTGCGCACTCC 1140
QY 1141 atcgatatccgaacactctcgtgtgctgacagcgtctggtgtgtgagcctgaggttatac 1200
DB 1141 ATCGATATCCGCAACCTTCCGTGCTGACGCTCTGCTGTGAGCTTGAGCTTGATTAAC 1200
QY 1201 actgacaaagttcaatgcagctcgtgtgatatgaagctccttgatgcgcgcgaactctgag 1260
DB 1201 ACTGACAAGGTCATGCGATTCGCTGCTGATGAAGTCTTGAGTGCAGCACTTCTGAG 1260
QY 1261 gtgcctgccaaactccagatgtctctcccgagagcggaactgagaccccgcaaac 1320
DB 1261 GTGCTTCCCAACCTCCAGATGTTCTTCCCGAGGCGCGCAACTGGAGCCCGCAAC 1320
QY 1321 cccactgatacagagacttcaactcgcgcgtgactaagagagtgagaaatcaacagga 1380
DB 1321 CCGACTATACGAGACTTTCACCTTGCGCGTGTCTATGAGAGTGAGAAATCAAGGA 1380
QY 1381 gtaacctctcgaatgtcgcgaacacgctcgtccgcaatgtgcgcgcgcgaactgttga 1440
DB 1381 GTTACTTTCGAGATGTGCGAAGACCGTCTGCGCAATGAGCCCGCACTGTTGAG 1440
QY 1441 atctgagcacttgagaacacactccaaagcttgagactaaccctgttcaactacactgt 1500
DB 1441 ATCTGGGAGCTTGAACCACTCCAAAGGTTGAGCTCACTCCCTGTTCACTTCACTCTGT 1500
QY 1501 gactccgaatcctctctgtcttccactgcgcgtgagagtgagacttaagagctgcgt 1560
DB 1501 GACTTCCGAGTCTTCTCTGTTCCACTGCGCGGTGAGATCGAGCTTATGAGGCTGCTGCT 1560
QY 1561 ctcaagagatgtctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1620
DB 1561 CTCAAGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1621 ccttcccgtaagctcgcgccttcttcaactaactggttcttcaactaactgactaact 1680
DB 1621 CTTTCCCGTAAGTCTGCGCTTTTACCTTAACGTTTACCTTACCTTACCTTACCTTAC 1680
QY 1681 gtggtgtcatatgtgtgacttgcacacacttgatccacagagagcagagagatagtgct 1740
DB 1681 GTGGTGTCTCATGTGTGACTTGCACACACTGATCCAGAGAGACAGAGATGATGCTG 1740
QY 1741 ctttcaatgtcactgtctcgtgtgactatgtctacaactacacagagttcatgtgaccca 1800
DB 1741 CTTTCAATGTACAGTGTCTGCGGTGACTATGAGCAACATCAACAGGATTCATTGACCCCA 1800
QY 1801 tggagcctctctctgagagcgcgcgcctctcctctcgtgagaggttcgagatgtcgtggt 1860
DB 1801 TGGAGCTCTCTCTGAGAGCGCGCGCCCTTCTCTCTGAGAGGTTGAGAAAGCTCGGCTG 1860
QY 1861 acttcaagcagcgttgcacatacactgaacgcgaattcagagatgagcttagcttaacccctag 1920
DB 1861 ACTTCAAGCAGCTTGGCATATCACTGACCGCAATTCAAGAGATGGCTAGCTTCAACCCCTAG 1920
QY 1921 ccaagcgtatgatatgacgctgagagtagaacggt 1958

QY 1383 taccttcgagatgctcgaacacgctctgctccgaatgtgccccgagacactgttagat 1442
|||||
Db 1447 TACCTTCGAGATGTCGAGAACCGCTCTCCGCAATGTGCCCGGACACTGTTGAAT 1506
QY 1443 ctggcgacttgagaacaaccccaacggttgactaccctgttcaattcacctgttga 1502
|||||
Db 1507 CTGGCGACTTGAACAACCTCCAAACGGTTGGAGCTACCCCTGTTCAATTCACCTGTTGA 1566
QY 1503 ctccgagctcttcctgcttccactgctgcgctgaggtgagcttataggtctgtgct 1562
|||||
Db 1567 CTTCGAGTCTCTTCTCTCTCACCTGCCGTTGAGAGTCGAGCCTTAAVAGGCTGCTGCT 1626
QY 1563 caagatgtgtctgctgcgctgctgctgaggtgtctatgttggagcccaactacgctcc 1622
|||||
Db 1627 CAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686
QY 1623 ttcccgtaagttctgcgcttllactaactcgtgttllactaactgttcaactcaact 1682
|||||
Db 1687 TTTCCCGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1746
QY 1683 ggttctacatgtgtgcctgcacacactgtatccacgagagcccaacatgtagtct 1742
|||||
Db 1747 GGTGCTACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1806
QY 1743 ttcaatgtcaatgttctgcgtgactatgtgtcaactacacgaggttcaactgaacct 1802
|||||
Db 1807 TTCAATGTCACTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1866
QY 1803 gagctctctgagagccccgccttctctcctcctcctcctcctcctcctcctcct 1862
|||||
Db 1867 GAGCTCTCTGAGAGGCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1926
QY 1863 ttccgagagctgtgcatactacacgagcttcaagagatgagcttagttcaacctacgct 1922
|||||
Db 1927 TTCAAGCGAGCTTGCATGACGACGCTGACGAGATGCTGACGCTTCAACCCCTACGCC 1986
QY 1923 caagctatgtagtgcgctcctcctcctcctcctcctcctcctcctcctcctcctcct 1953
|||||
Db 1987 CAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2017
RESULT 6
AX108674 2095 bp DNA Linear PAT 30-APR-2001
LOCUS AX108674
DEFINITION Sequence 3 from Patent WO0121748.
ACCESSION AX108674
VERSION AX108674.1 GI:13923901
KEYWORDS
SOURCE Stachybotrys chartarum.
ORGANISM Stachybotrys chartarum.
REFERENCE 1 (bases 1 to 2095)
AUTHORS Convents,D.V., Doornik,M.U., de Vries,C.H. and Wang,H.
TITLE Detergent compositions comprising phenol oxidizing enzymes
JOURNAL Patent: WO 0121748-A 3 29-MAR-2001;
FEATURES
Location/Qualifiers
source
/organism="Stachybotrys chartarum"
BASE COUNT 437 a 618 c 510 g 530 t
ORIGIN
Query Match 99.4%; Score 1946.2; DB 6; Length 2095;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 127 CGGAGCTCTGTGATGTCAGATCCGTTGCTGTGATGATGAGACATGCTCCGGTCT 186
QY 123 caacgaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 182
Db 187 CACCAAGAGGCGAGAGCGAGCTAGTCCCTCCCTGCTTGTGAGAGGCTCTCCGAT 246
QY 183 cccctccttgaaagggcccaagtagtaagtatactcttaagcttagcagagcagcagc 242
Db 247 CCTCTCTTGAAGGGGCCCAAGTAAATAGTACATTCTATAGCTTAGCAGAGCCACGTTG 306
QY 243 cttaactatgtagtgcgttcccaaccccaacacactcgtgagagagacatctgtactacgag 302
Db 307 CTAACTATGAGAGTACCGGTCGCCCAACCCCAACACTGAGAGAGACATCTTGTACTGAGA 366
QY 303 tggagataagcagcctcttcccaacagctacacagctcgtatcgtgagcagcagcagatgtgt 362
Db 367 TGGAGATTAGGCCCTTCTCCACAGATCTACCTGATCTGGAGCCGGCAACATGTTG 426
QY 363 gatacagatgcaatgtcccaagagctacacatcatgcttccctgtagcagctgaggtgtg 422
Db 427 GATACGATGGCATGTGCCCAAGACCTTACCATCATGTCTCTGTGACATGAGAGTGTG 486
QY 423 tccgcttctgtaacagcagagagacacactctcccaacagcgttccacttgacagcgtct 482
Db 487 TCCTGCTGTGAACAGCGAGAGAACACCTCTCCCAACAGCGCTCCACTTGCACGGCTCTT 546
QY 483 tctctcagctccctctgtatgtgtgtgagtagacatacccaagcctggtcgagtagcaag 542
Db 547 TCTCTGAGCTCCCTTGTATGTTGGCTGAGGACATACCCAGCTGCGAGTCAAGG 606
QY 543 attactactaccccaacagcagcgtcccgatgcttctgtgacatlaacatgacatg 602
Db 607 ATTACTTACTACCCCAACAGCGAGCTGCGCATCTTGTGACATACCATGTCATGCTAGT 666
QY 603 ccatcacccgcagagagacgactacatcagtgagtcaggtgcttactatgtccagagccgg 662
Db 667 CCATCACCGCGAGAACGCTATCATGAGGTGAGGTGTGTACTATGATGATGATGATGATGAT 726
QY 663 ctgagagatgctcctgaactcccccagcgtcagcagcttgaatccccctgtgtctga 722
Db 727 CTGAGGATGCCCTGAACCTCCACGCGCTGAGGCGAATTTGATATCCCTTGTGTTCTGA 786
QY 723 ctgccaagcgatatacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 782
Db 787 CTGCCAAGCGATACAAACGACAGCGACTCTTCTTCCACCAATGAGAGGTTTCCAGCT 846
QY 783 tctgggtgagcgttactcaagtgtagtgaagttgagccatltgagatgtctcagatccatgaa 842
Db 847 TCTGGGATGAGTATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 906
QY 843 gtagcatgtatgaattgtgcatgctcttaacagtgctatcatcagaaacggtcagccttg 902
Db 907 GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 966
QY 903 gctatgttcaacgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 962
Db 967 GCTATGCTCAACGTGACGCGCGGCAAGTACCGCTTCCGCTTCCCAACGCTGCGCTGTC 1026
QY 963 agcagcttctcgtatctcttgaactctgaagatctcagagacagagacagcagcagcagc 1022
Db 1027 AGCGCTTTCGCTCTGTATCTTGTACTCTGAGATTCAGAGACCAACATCTTCCCTTCCA 1086
QY 1023 ggtcaatgcagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1082
Db 1087 GGTCAATGCCCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
QY 1083 ctctatgacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1142
Db 1147 CTCTATGGCCGAGCGCTGAGGCTGATGACATCTTCCACCTTCCGCTGCGCAATGCAT 1206
QY 1143 cgaatccgcaacactctcgtgctgagcagcagcagcagcagcagcagcagcagcagcagcagc 1202

Db 531 AAGACTATTCTACCCAAATGACAGTCTGCTGATCCCTATGCTATCCAGATCATGCT 590
 Oy atgtccatccagccgagaaacgctacatgggtcagctggtgtctacatgacccagac 658
 Db 591 ATGCATATCACTGGTGAAGAACGGCTACCGTGGCCAGGTGGTGTCTCTACATGCTCATCTAC 650
 Oy cccggtgagatggccctgaaccccccagcggtacggcgaattgatatcccttggt 718
 Db 651 CCAGCCGAAGACGGCTCTCAACTTGGCAAGTATGCGAGTTGCGATTCATTCATCATGATC 710
 Oy ctgctgccaagcgataaagcagcagcagcctctctccaccaatggagaggttcc 778
 Db 711 CTCAAGTCCACATATATACGCAAACTGTCACCATTAAGAGAGCTGTAAC 770
 Oy agctctggagtgagcgtattcctaagtggtaagttgagccattgagatgctcaagct 838
 Db 771 TCATTCTGGGGTGAATGTAATTCACGT 796
 Oy 839 agaaatcagatgataaattgtgcaatgctcacaacagtgatcacaagaagctcagc 898
 Db 797 -----GAACGGTCAAC 807
 Oy 899 ctggcctaigtcaagctgacgcgcgcaagtaacgcttcgcttccataacgctgcg 938
 Db 808 CCGGCGCTTTCAGACGCTTGAAGCTCGCAATATCGATTCGGTCCCTCATGCGCGAG 867
 Oy tctacgctcttcgctctgctgctgctgctgctgctgctgctgctgctgctgctgct 1018
 Db 868 TTTTCTCGCTCTTTGGGCTTACTTGTGATGATGATGATGATGATGATGATGATGATGATG 927
 Oy 1019 tccagtgatctgcgctgacgtgctgctgctgctgctgctgctgctgctgctgctgct 1078
 Db 928 TCAGGATATATGCTCGCATCTGCTGCTTCTTGAACACCTCGCGATTCACGCTTGGCT 987
 Oy 1079 acatctctatggcgcgagcgtgagaggtgtgtatcagatcttccactcgtcgtgcag 1138
 Db 988 ACATTTCCATGCGCGAGCGTATGCAAGTGTGTTGACTTCCGACTATCTCCGACCAAG 1047
 Oy 1139 ccatacgtatccgaacccctcgtgtg---ctgaagctcgtgctgctgctgctgctgctgct 1195
 Db 1048 CTATTTGAATCTCGCACTGGCGGTAGCATTTGGCGGCGATCGGAACAGATACCGACTATG 1107
 Oy 1196 ataactgacaaggtcagatcgtcgtgtgtatgaagtcctgtgctgctgctgctgctgct 1255
 Db 1108 ACAACACGCAAGGATGATGCTTGTGTAAGCAGACACAACTCAGCCAGATACCT 1167
 Oy 1256 ctgaggtgctgcgcaacccctcagatgtctctcccgagggcgcaactggagaccccg 1315
 Db 1168 CAGTTGTTCTGCTAACCTTGTGATGCTTCCCTTCCTCCACCAACCAACACCCCTCC 1227
 Oy 1316 caaaccccaatgatacgaagcttccactcgcgcgtgtcctaagtgagagagcaatca 1375
 Db 1228 GACA-----GTTCCGCTTGGTGGCGACCGGCTCTCACTGGACTATTA 1269
 Oy 1376 accgagatcactctcgaatgctcgaagacgctcgtcctcgaatgctcgcgcgcgacactg 1435
 Db 1270 ATGCTGTGCTTGTGATGATGATCAAAACGCTGCTGTAAGAGTACCCCTGCTGATG 1329
 Oy 1436 ttgagatcgtgcgacatgagagacaactccaaggtgtgactcaccctgttcaattcaac 1495
 Db 1330 TCGAGCTTGGAGCTATCAACCGCGTACGGTGTGACACCCCTATTCATTCATTCATC 1389
 Oy 1496 tctgtaactcgcagctcctcctcgtc-----cagcgccggtgagtgagcctt 1546
 Db 1390 TTGTGCTGACTTAAGTGTATTTCTGCTACTTCGGCAACACAGCGCGACACTCATGCCAT 1449
 Oy 1547 atgagctgctgctgctcgaagatgtgtctgctgctgctgctgctgctgctgctgctgctg 1606
 Db 1450 ACGAG-----TCGGTCTCAAAACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506
 Oy 1607 aggcacactaagctcctcctcccgtaagttctcgccttactaactaactggtttcaactat 1666
 Db 1507 AGGCTCATTAAGCGCGCTTTCC----- 1528

Oy 1667 gctaacatcacaagtggtgtctacatgttgactgcccacaacatgataccagagacca 1726
 Db 1529 -----TGGTGTATATATGTTTCATTTCCACAAATTTGATTCACGAGATCA 1573
 Oy 1727 cgaatgatgctgctccttcaatgctcgtctcgtgctgctgctgctgctgctgctgctgct 1786
 Db 1574 CGATATGATGCTGCTTTCATACGCGACCGCTCGACATATATGCTATATGCTACTGT 1633
 Oy 1787 gttcatgaccccatgagagctcctcgtgagggcccgcccttctcctcgtgagagttcga 1846
 Db 1634 TTTGGTTGACCTTATGGAAGAGCTTTGGCAGGCTGCTCTTATGAACTCGCGAGATTCCA 1693
 Oy 1847 gaatgctgctgctgctcgaagcgtgctgctcactacagcagcagcagagagagctag 1906
 Db 1694 GGCCTCAGAGTGGCCAGTTCACGCTTCACGCTTCATGCTATGAGCTATTCAGACTATGCTGA 1753
 Oy 1907 ctcaacccctcagccagcgtgatga 1933
 Db 1754 ATACAGACTTACGCCGCGAGCTGACGA 1780
 RESULT 8
 E05283
 LOCUS Billirubin oxidase gene. 1959 bp RNA linear PAT 29-SEP-1997
 DEFINITION E05283
 ACCESSION E05283
 VERSION E05283.1 GI:2173473
 KEYWORDS JP 1993199882-A/1.
 SOURCE Myrothecium verrucaria.
 ORGANISM Myrothecium verrucaria.
 Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Myrothecium.
 REFERENCE 1 (bases 1 to 1959)
 AUTHORS Ando, K., Koike, S., and Samejima, T.
 TITLE PRODUCTION OF BILLIRUBIN OXIDASE
 JOURNAL Patent: JP 1993199882-A 1 10-AUG-1993;
 AMANO PHARMACEUT CO LTD
 COMMENT OS Myrothecium verrucaria
 PN JP 1993199882-A/1
 PD 10-AUG-1993
 PF 24-JAN-1992 JP 1992034126
 PI ANDO KEIICHI, KOIKE SATORU, SAMEJIMA TATSUYA PC
 C12N1/53, C12N1/19, C12N9/06, C12N15/53, C12R1.645; CC
 strandedness: Double;
 CC topology: Linear;
 CC *source: Strain=MF-1;
 FH key Location/Qualifiers
 FT CDS 66..1784
 FT sig_peptide 66..179
 FT mat_peptide 180..1781
 FT Location/Qualifiers
 FEATURES
 source I. 1959
 /organism="Myrothecium verrucaria"
 /db_xref="taxon:5532"
 BASE COUNT 457 a 549 c 436 g 517 t
 ORIGIN
 Query Match 27.0%; Score 528.6; DB 6; Length 1959;
 Best Local Similarity 62.0%; Pred. No. 1,5e-114;
 Matches 1058; Conservative 0; Mismatches 474; Indels 175; Gaps 7;
 Oy 239 gtgtctatcatgtcagttacgctcccaacccaacacatgagagagatctgtactac 298
 Db 237 GTTAAGCAGCCCGCTTGACTGTAACTCTGTAATGAGCAAGATCTGCTACTAT 296
 Oy 299 gagatgagatgagcccttcccaacagatctacacatgctgctgagcgccgaacatg 358
 Db 297 GAGGTGAGATCAAGCCCTTACACACAGGATGATACCTGACCTTGATGATCGCTGATCTG 356
 Oy 359 gtgtatcagatgagcagctccagacatcactatcgtctcgtgagcagagat 418


```
CDS
98..1906
/gene="ppoa"
/codon_start=1
/product="polyphenol oxidase"
/protein_id="CAB75422.1"
/db_xref="GI:696278"
/db_xref="SPTREMBL:Q9P8C3"
/translation="MVATPALRALALVLSLKAQAMPKPELDIPEEAAALAIAYED
PANDLORSPANDLORSPAYLFPQAPISIPVKEPLFTVTPYNGEIDYEIE
IKHSEQEPDLPADLYGDISPGPFIPEKGESEYRVNKAIVSSYLHSPS
RAPDGMADVNTPEEKDYYPNAQARFSYHDHAAHETAEANYNCOAGYTIHDP
AEDSLGPSGEYDPLILTSKOYNSDGLSTKGTSLMGDVIQVNGVWPVEDY
EPKRYRFLDLAAVRSFSLYEVDPADDTRIPEQVIAVSDSLPEPTTSLVISA
ERYEILDFDSEKTEILRNEPAVGGLIEVNDTDKVMFNVAEGPLSPDSYV
PSTLRDVPSPSTSTTIDHSPFAPACOMSNVTFSDVBRILLANVPLIGVOLMOT
TMAAGRTHPHIHVDKILSRGAATGATRTREYEPESAGLRVYLKGETVVE
AHYAPGVYMFHCHNLTHEDDMAATNVLYPDGYNSPALADPMODEFRAPKYVD
NDVEVERENAFSTDEIEAOVDMASYLPYDNDPMEIV"
161..280
/misc_feature
/gene="ppoa"
/feature="polyphenol oxidase propeptide"
281..1903
/mat_peptide
/gene="ppoa"
/evidence="experimental"
/product="polyphenol oxidase"
1907..2120
/3'UTR
/gene="ppoa"
2073
/polyA_site
/gene="ppoa"

BASE COUNT 442 a 723 c 546 g 415 t
ORIGIN
```

```
Query Match 26.4% Score 517.4 DB 8: Length 2126:
Best Local Similarity 61.4% Pred. No. 6.8e-112:
Matches 1041: Conservative 0: Mismatches 476: Indels 178: Gaps: 6:
```

```
QY 797 attcaagtgtaagtgagcccatltaga1tctcagatcctagaag1atcga1tga1ta 856
DB 887 ATCCAGGT----- 894
QY 857 aattgtcatgtcctacaacag1tgc1ta1caagaag1tcaagcct1tgcctatga1taag 916
DB 895 -----GAACGAGATCCCGTGGCCCTACTTGTGAGC 923
QY 917 tgaagcgcgcaag1taccgt1tccgt1tcc1ta1aagc1tccgt1tcaagcct1tgc1tc 976
DB 924 TTGAGCCCAAGAGATACCGCTTCGCGCTGTCAGCCCGCGCTGTCTGCTCTTCTTCGC 993
QY 977 tga1tct1tca1cct1tga1gga1t1cga1ga1ccag1tcc1tcc1tcaag1tca1tgc1tc 1036
DB 984 TCTACTTTGCGACACCGCGGATGAGAGACCCGATCCCGTTCCAGATCATCGCTCCG 1043
QY 1037 acg1tgc1tgc1tga1gg1cc1t1tga1ca1tga1ca1c1tgc1tga1c1tca1tgc1tgc 1096
DB 1044 ACTGTGCTCCTCGAGAGCCCGTCACGACCTCAAGCTCGTCTCTCATTTGCTGAGC 1103
QY 1097 gctggag1t1gt1tca1cct1tcc1tcc1tcc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1156
DB 1104 GCTACGAGATCATCTCGACTTCTCGACTTGAAGGCAAGACCACTGAGCTCCGCAACG 1163
QY 1157 ttc1tgc1tgc1tga1cgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1216
DB 1164 AGCCCGCGTGGCGGCGCTCGCGATGAGTCAACTACGACGACGACCGAACAGTCAATGC 1223
QY 1217 gatt---cglc1t1tga1tga1tcc1t1tga1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1273
DB 1224 GCTTCAACGTCGCGAGGCGCCCTCTCTCGCCGACACCTCGTCTGCTCCACCC 1283
QY 1274 tccga1tgc1tcc1tcc1tcc1tcc1tcc1tcc1tcc1tcc1tcc1tcc1tcc1tcc1tcc 1333
DB 1284 TCGTGAACGTCCTTCCC-----TCCAGACCTTGACACCATGAGACC 1328
QY 1334 agact1tca1cct1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1393
DB 1329 ACTGTTCCGCTTCGCGCGACCGCGCGCGGCGGAGGACATCAACGAGTCACTTCTCCG 1388
QY 1394 atgtcga1ga1cgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1453
DB 1389 ACTGTCAAGAACCCCTGCTGCGGCAACGTCGCTTGAGACCGTGTGACCTGA 1448
QY 1454 agaca1cct1tca1cgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1513
DB 1449 CGAACCGCGCGAGGCGTGAACGATCTATTCAATTCATCTTGTGACTTAAGATTTC 1508
QY 1514 tttctc-----gt1tca1cgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1555
DB 1509 TGTCTGCAACGGGTGCTGCTAGAGGATGCTACGACCGGTGTGTCGAGCGTACGAGTCTG 1568
QY 1556 ctgtctca1ga1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1615
DB 1569 CTGCTGTGAAGAGAGTCTGCTACCTCGGCAAGGAGAGAGTGTGTCGTGAGAGCTCACT 1628
QY 1616 agc1tct1tcc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1675
DB 1629 ACGCTCTTGCCCC----- 1642
QY 1676 taca1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1735
DB 1643 -----GGGTGTACATGTTCCATGCGCAACACCTCATCAGAGAGAACGACATGAT 1695
QY 1736 gctgtcttca1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1795
DB 1696 GGCGGCTTCAACAGTCAACCTCTCCCGACTACGGCTTACACTCGACGGCCCTCGCTGA 1755
QY 1796 cccat1tga1cct1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1855
DB 1756 CCCATGCAAGAGAGATTCGCGCCCAAGCCGTACGTGAGACAGAGCTCAAGGTGCCGA 1815
QY 1856 ggt1tga1tca1cgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc1tgc 1915
```


Db 299 GCACGGGGGGCCCGTGGCGATCCTACACCAACGGGCTCGACACCCACCGCAGGTGC 358
Qy 468 acttgagagcgtcttctcgtgagctcccttgatgttggtgaggaactaccagc 527
Db 359 ACCTGCACGGGGGACAGTCCGGCCACCGACAGGTACCCGATGGACTATCCCGC 418
Qy 528 ctggcagatcaagaatctactaccccaagcagagcgtgcgcgcatgttgatcc 587
Db 419 CGGGCGGCTCGAAGGCTACGACTACCCGAACTTCAGCGCGGCGGACGCTCTGGTACC 478
Qy 588 atgacacatgcattcattaccccgagagagcctacatgggttcagagctgtgtctaca 647
Db 479 ACACACACACCCACGGCTACAGGCGGACACCTTACCGGACCTGACGCGCTTCTATC 538
Qy 648 tgaatcaagaccccgctgagagatgccttgaaactcccgagcgtctacggttgatata 707
Db 539 TGATTCACAGACCGCGGCGAGCATCCTGCGCTGCGCGCGGCAAGTACGAGTGCCTGA 598
Qy 708 tcccttggcttctgactgcgaagcgaatacagcagagcagcactctcttcacccaatg 767
Db 599 T--CATGCTGGCGCAACGCCCACTGACGACTCCGGCGGCTCTGCTTGGCGCACCGCG 655
Qy 768 gagaaggttccagactctctgggtgagc 794
Db 656 ACGACCGGCTCACCATCTCGGCGAAGC 682

RESULT 12
AX195929 109519 bp DNA linear PAT 28-AUG-2001
LOCUS AX195929
DEFINITION Sequence 1 from Patent WO0151639.
ACCESSION AX195929
VERSION AX195929.1 GI:15386161
KEYWORDS
SOURCE
ORGANISM
Micromonospora carbonacea.
Micromonospora carbonacea.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Micromonosporineae; Micromonosporaceae;
Micromonospora.
1 (bases 1 to 109519)
Hosted, T.J., Horan, A.C. and Wang, T.X.
Evernimomichin biosynthetic genes
Patent: WO 0151639-A 1 19-JUL-2001;
Scherling Corporation (US)
FEATURES
source
1. 109519
/organism="Micromonospora carbonacea"
/db_xref="taxon:47853"
BASE COUNT 15101 a 39910 c 39406 g 15102 t
ORIGIN

Query Match 4.9%; Score 95.8; DB 6; Length 109519;
Best Local Similarity 54.8%; Pred. No. 5.2e-12;
Matches 212; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

Qy 408 gcaatgagatgtgtgcgtctcgtgaacagcagagagaacacactctcccaacagcgtc 467
Db 23046 GCACGGGGCGGCGCGGCGGATGATCACTACACCAAGGGCTCGACACCGCAAGAGTGC 23105
Qy 468 acttgagagcgtctctcgtgagctcccttgatgttggttggtgaggaactaccagc 527
Db 23106 ACCTGCACGGGGGACAGTCCGGCCACCGACGAGCTTACCGGATGATCCCGC 23165
Qy 528 ctggcagatcaagaatctactaccccaagcagagcgtgcgcgcatgttgatcc 587
Db 23166 CGGGCGGCTCGAAGGCTACGACTACCCGAACTTACGCGCGGCGGACGCTTGGTACC 23225
Qy 588 atgacacatgcattcattaccccgagagagcctacatgggttcagagcgtgtgtctaca 647
Db 23226 ACGACACACACCGAGCTACAGAGCGGACGACTTACCGCGGACGAGCTTCTATC 23285
Qy 648 tgatccagagaccgctgagatgtccctgaacctcccgagcgtctacgagcgttgata 707

Db 23286 TGATGCAGACCGCGCGGACATCATCTGCGCTGCCCGCGGCAAGTACAGCTGCCGA 23345
Qy 708 tcccttggcttctgactgcgaagcgaatacagcagagcagcactctcttcaccaatg 767
Db 23346 T--CATGCTGGCGCAACGCCCACTTGCAGCAGTCCGGCGGCTCTGCTTGGCGCACCGCG 23402
Qy 768 gagaaggttccagactctctgggtgagc 794
Db 23403 ACGACCGGCTCACCATCTCGCGAAGC. 23429

RESULT 13
AX204969/c 3446 bp DNA linear PAT 30-AUG-2001
LOCUS AX204969/c
DEFINITION Sequence 4 from Patent WO015180.
ACCESSION AX204969
VERSION AX204969.1 GI:15394251
KEYWORDS
SOURCE
ORGANISM
Micromonospora carbonacea.
Micromonospora carbonacea.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Micromonosporineae; Micromonosporaceae;
Micromonospora.
1 (bases 1 to 3446)
Staffa, A., Zazopoulos, E., Mercure, S. and Nowacki, P.
Genetic locus for evernimomichin biosynthesis
Patent: WO 015180-A 4 02-AUG-2001;
Ecopia Biosciences Inc. (CA); Farnet, Chris (CA)
FEATURES
source
1. 3446
/organism="Micromonospora carbonacea"
/db_xref="taxon:47853"
misc_feature
3..1037
/note="ORF 2 (positive strandedness) incomplete;
C-terminus only (N-terminus undetermined)"
misc_feature
1077..2231
/note="ORF 3 (positive strandedness)"
misc_feature
2242..3444
/note="ORF 4 (negative strandedness) incomplete;
C-terminus only (N-terminus undetermined)"
BASE COUNT 427 a 1148 c 1315 g 556 t
ORIGIN

Query Match 4.2%; Score 82.8; DB 6; Length 3446;
Best Local Similarity 55.1%; Pred. No. 5.9e-09;
Matches 162; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Qy 422 gtccgctcgtgacacagcagagagaacacactctcccaacagcgttcactgagcgtct 481
Db 3351 GTCCGATACACTACACCAAGGGCTGAGACAGCACGAGACGAGTGCACCGGTGGC 3292
Qy 482 ttcttcgagatcccttgatgttggttgaggaactaccacagcgttgagatcaag 541
Db 3291 CGGTCGCGGCGAGCAGTGCAGTGCAGCCCGATGATATCAGCGGGCGGCTGCGCG 3232
Qy 542 gattactactaccaccaagcagcagcgtgcgcgcatgttggtgacatgacatgcacg 601
Db 3231 GTTTAGACTATCCCGAAGCTCCACCGCGGGCCACGCTGTGTACACGACGACACCCAC 3172
Qy 602 tccatcaccccgagagacgctactatgttgatgagcgtgtgtctacatgtaccagccg 661
Db 3171 GACTACGAGGCGGACACAGTCTACGCGGCTGACAGGGCTTCACTCACTCAAGAGACCC 3112
Qy 662 gctgagatgacctgaacctcccgagcgtctacgagcgttgatcccttg 715
Db 3111 GCCGAGCGGACCTCCACCTGCCCAAGGGCGGCTACGAGCTCCCGATCTGCTG 3058

RESULT 14
AY033994/c 9531 bp DNA linear BCT 18-JUN-2001
LOCUS AY033994/c

DEFINITION	Streptomyces aureofaciens type-II polyketide gene cluster, partial sequence.	gene	complement(2510. .3367)
ACCESSION	AY033994	CDS	/gene="aur2g"
VERSION	AY033994.1		complement(2510. .3367)
KEYWORDS	GI:14486270		/gene="aur2g"
SOURCE	Streptomyces aureofaciens.		/note="Aur2g"
ORGANISM	Streptomyces aureofaciens		/codon_start=1
REFERENCE	Bacteria: Firmicutes; Actinobacteria: Actinobacteridae; Actinomycetales; Streptomycetaceae; Streptomyces. 1 (bases 1 to 9531)		/transl_table=11
AUTHORS	Kormaneec,J., Bistakova,J., Novakova,R., Homeroova,D. and Rezuchova,B.		/product="oxygenase-like protein"
TITLE	Cloning and characterization of a new polyketide gene cluster in Streptomyces aureofaciens CCM3239		/protein_id="AAK61715.1"
JOURNAL	Unpublished		/db_xref="GI:14486274"
REFERENCE	2 (bases 1 to 9531)	gene	/transl_table=11
AUTHORS	Kormaneec,J.		/product="cyclo-oxygenase-like protein"
TITLE	Direct Submission	CDS	/protein_id="AAK61716.1"
JOURNAL	Submitted (02-MAY-2001) Gene Expression, Institute of Molecular Biology, Slovak Academy of Sciences, Dubravská 21, Bratislava 84251, Slovak Republic		/db_xref="GI:14486275"
FEATURES	Location/Qualifiers		/translation="MTPGQREVEHEITVSAPAAVYRLIAEVNMPRIPTIYVDO VPBGRGERIRIATANGCAKNRTSRTLDALRIETFRQEVTPAPPAAGCTIIIE LSAOESVRRLHDYRAVDADDLSLAWIDEADNRSSSELAALKNAEAAHAEELITS FEDVQVIEGSAKDDYDFVNEADRWHEHLPHVAVRLREDDPGLQELMDIAKAGSVH TTKSYRVTFEAFHRIAYAKOTLLPALMTLHTGTWFEENEEDGVAASSOHTVIINDNIAR ILGSAIVEDAKRKYHTALSTNSATLTHAKAAYAEORAGR"
source	1. 9531	gene	complement(4349. .5134)
	/organism="Streptomyces aureofaciens"		/gene="aur2e"
	/strain="CCM3239"	CDS	complement(4349. .5134)
	/db_xref="taxon:1894"		/note="Aur2e"
	<1. 212		/note="Aur2e"
gene	/gene="aur2j"		/codon_start=1
CDS	<1. 212		/transl_table=11
	/note="Aur2j"	gene	/product="ketoreductase-like protein"
	/codon_start=3		/protein_id="AAK61717.1"
	/transl_table=11		/db_xref="GI:14486276"
	/product="aminotransferase-like protein"		/translation="MSDQGRVAVVGTSGIGLAVARLLAVOGHRYFTGARSANVA ETYKELOGEGLVDVAVDVRSTGVAMIOAADVRGSVDVYVNNAGRSQGVYADI ABEELMHVIDTNNISVFRREALFTGLIRAKDGRGIINIASTAGKGVYAPYSAS KHGVCFETKALGNELAPGTFVAVNVCQCYETPRNAQVROCYAAYATREAILLEKQ AKTPLGRISTPEEYAGLVGLASDAASITISQALNVGGJGNF"
gene	complement(345. .1772)		5386. .6657
CDS	/gene="aur2i"		/gene="aur2a"
	complement(345. .1772)		5386. .6657
	/note="Aur2i: similar to Bacillus subtilis spore-coat protein CoCa"	gene	/note="Aur2a"
	/codon_start=1		/codon_start=1
	/transl_table=11		/transl_table=11
	/product="oxidoreductase-like protein"		/product="B-ketoacyl-ACP synthase-like protein"
	/protein_id="AAK61713.1"		/protein_id="AAK61718.1"
	/db_xref="GI:14486272"	CDS	/db_xref="GI:14486277"
	/translation="MAAAGGAALFPMDRITLGASQAVASSLATAEAPIVPFAHAMP LPKNIKPTSTFATSDLYEIRMOEAQVEIVKIGMSKVITGTPGPIRTATOGREVVV ROINELQVNTAVHLHGAHVLSHDGLPMQIVRGERTYRPNQPAASJLMDHHAH LEANVFMGLHGLILTDHNEKRLDLPSPGPDVPLVIRDAVEADGTLITRPSDCPH MLVNGKERPEYQAAKRYRFRVYNACANRYKLFADGIEFTQIGTGGLFEPVQOS ELLMLGERADIVDEFSRYKVDVSVLENQAOISIERPAMREDIVRTADYSFVRP LTTYPQPTVEDEFEIRTFEPAMTINGOSIDPWRVDITAKLGTTEVATRVNVAAPA BGDPDFHMSHFHLLTYERVLBNKCPACTRDLGLKDYTLGPGDVIKIAMTMGPPT GQYLHCHQHLGSSGGMGRIDIVA"		654. .7916
gene	complement(1782. .2513)		/gene="aur2b"
CDS	/gene="aur2h"		654. .7916
	/note="Aur2h"		/gene="aur2b"
	/codon_start=1		/note="Aur2b"
	/transl_table=11	gene	/codon_start=1
	/product="unknown"		/transl_table=11
	/protein_id="AAK61714.1"	CDS	/product="chain length factor-like protein"
	/db_xref="GI:14486273"		
	/translation="MSVAREIVAAAGYRRAGCTSGNGPGVAGREBDDPGGAPGA VPGVAGRGAGRGALMLISSGVPIGLLAFQAFYANIVYTRTAPPPQEDIHAW WHSGVPEPEAFRPGFLALVAFTVLTPLLALAGVLRRCGLVAGAGRSIPAL GFAAAATVAGAVVTVPLVILARAFSPVDGSGPLGAVVYADAVTAMRLCALALLAFVA GAWMAAGAPSLRRESGSLAEBS"		

misc_feature
MAGPLIVCGRCAGPKARNATIVLPHEKSLYLTPOLSSSSRCYRHHRLILSLKYVDERR
ROHIVCGCRRRRLRLSSQSSSSQPPANVYTPRRPPSNAPRRIRIGLKSTPPTPQQNPDAHLAA
EAVAS4"

gene
complement(4956..8529)
/gene="P0037C04.2"
complement(4956..8529)
/gene="P0037C04.2"
/note="probably inactive due to frameshift(s) in CDS
probably inactive due to stop codon(s) in CDS
pseudogene, similar to gag.pol polyprotein"
/pseudo
join(10530..10652,12160..12528)
/gene="P0037C04.3"
join(10530..10652,12160..12528)
/gene="P0037C04.3"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB5516.1"
/db_xref="GI:14090358"
/translation="MDIPHNEERLLDYREEKSRDQDKLMEVGSMAATDKNTKPNNN
LSTIVTYLVHLYREDSRVFLNLNGSDVPRGRSLSYVCVAAGAEALYTMASGAP
GALSTASGSPVPRVRRWRYPTGGLFVPDAVWRASGSVSIVHWKLLPPELEEDT
TAA"

gene
complement(15496..16914)
/gene="P0037C04.4"
complement(15496..16914)
/gene="P0037C04.4"
/note="contains EST C28475.C61239)
similar to Oryza sativa chromosome 1, P06980A4.16
unknown protein"
/codon_start=1
/protein_id="BAB5517.1"
/db_xref="GI:14090359"
/translation="MADRPSPCVLIERVRFVEAGLTSGASRDADVAATIEAGSM
SRSYOMMASAEEMERRMAPSVKPAFLADPOASLHMLIPAPRTLLICISGT
HGIYVYVIAHRCYTLIDASNHLTAIPYDPDSVYPLGRSAVLVSAGADDDITLAD
IYTSQRSRINPALPKRAITFAVKNKGWIOSLTHPLPLPHLCGPYFHHIDAFSEF
ACTITMWDILKGLICDELSSPDGRPLVPLRHCIDAHDPKRCFSPDGRHSIGRV
SAIFALVIGICEASCPANEVCLKTWSLSPDEHKMEETTLVGDIMASESNOMGSG
LHVHPEFSLIGVEDGIMYALVNHKEEPIRLNPEGDSIGIOLVPRKANMIRFDMIL
QKVLSTSLSKKPPFRNFTMELASDSATLDQROYNLITFTLITLHTTSLADHP
NMLYCFTLFQIOLTLIFVTS"

gene
complement(join(18962..19079,19226..19636,19997..20204),
20474..20575,20871..20948,21217..21321,21330..*21552)
/gene="P0037C04.5"
complement(join(18962..19079,19226..19636,19997..20204),
20474..20575,20871..20948,21217..21321,21330..21552))
/gene="P0037C04.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB5518.1"
/db_xref="GI:14090360"
/translation="VVALVGGEDVEDATMGWCHRRPPIRLCGCGRQGGYRRPDP
VAGGVPEKASRSRGIPDDPMPDPDAVVGGGGGRIRIRLRPLSGEGGGGGWRSRA
EETQSSGDEKREERKRREDEEERREKRLIYVEVLSDEMLITNPSOMGSGVVA
YLKTRTSDSDRSVAQKODPTEASDSDHVDGGRAPSNHAPRTETIMPTPLIAS
DVARNARYOMIEPVGAHILGRLLIPVARNLLPVAASRRPDPADALAAARIV
VVAAPFSSSSRRRLPSSPLSSPTPAALAFSLAGRRARLLSRRRPSSPLSSPAP
AAELFSLAGHAARRRRRSTPPRSPAPAVLSAASLSSCNOTLSKKRRPCITRY
PTNQIHEVRMSRVAAAEGR"

gene
join(25223..25413,25596..26220)
/gene="P0037C04.6"
join(25223..25413,25596..26220)
/gene="P0037C04.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB5519.1"
/db_xref="GI:14090361"
/translation="MRLTRIGMMKKNKGRRTTMDGTAKTRYSVAKKTEGCLTDOO
GVEEETTKMYTSDISDLACRPNNEPILGIRLCPLEKLYLAASAKCTPRASCKMT
VGFTEHLSEASCTHAPCGCPDCTGYIGLLYNIHIDEHATDVAAGSGRGTIVY
HKSKEPVNLHGGTGGNRFLNLNGDVLISGRSLVSGPPTANCELLYIELA
ADGPGPCVGLTILASGTVPCVRLRLEGGNAAFLEVPDSYSGSSDVTASVYL"

CDS
gene
CDS

[illegible]

Mon Oct 7 11:28:36.2002

us-09-656-640a-1.rge

Page 18

QY 665 GAGATGCGCCGAACCTCCCAAGGGGTAGGGAGTTTATAACCCCTGTTCGACT 724
| | | | | | | | | | | | | | | | | | | | | |
Db 131915 GTGAGAGGCCGCCTCGCCTCCCTGCGCGCATATAGTTCCACCOCCTCTCATCTGCTGCC 131974
| | | | | | | | | | | | | | | | | | | | | |

QY 725 GCCAAGCATACAAGCAGAAGAGCACTCTCT 755
| | | | | | | | | | | | | | | | | | | | | |
Db 131975 GACCGACCTTTTAGCCGACGAGCGCTCATCT 132005
| | | | | | | | | | | | | | | | | | | | | |

Search completed: October 5, 2002, 20:23:09
Job time: 8724 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 17:59:31 ; Search time 448.02 Seconds
(without alignments)
7503.501 Million cell updates/sec

Title: US-09-656-640A-1

Perfect score: 1958
Sequence: 1 gatacctacatgacatgacag.....ccgcctgagagtagaccggt 1958

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1958	100.0	1958	22	AAF30028
2	1956.4	99.9	1958	22	AAF82585
3	1946.2	99.4	2095	22	AAF82586
4	1946.2	99.4	2095	22	AAF82586
5	528.6	27.0	1959	14	AA047790
6	517.4	26.4	2110	21	AA261243
7	460.4	23.5	1791	20	AA227601
8	460.4	23.5	1791	21	AA550019
9	460.4	23.5	1791	21	AA551314

10	458.8	23.4	1791	20	AA225727	Stachybotrys chart
11	300.4	15.3	2067	20	AA225735	Stachybotrys chart
12	300.4	15.3	2067	20	AA227609	Stachybotrys pheno
13	300.4	15.3	3676	20	AA225728	Stachybotrys chart
14	300.4	15.3	3677	20	AA227602	Stachybotrys pheno
15	300.4	15.3	3677	21	AA550018	Stachybotrys chart
16	300.4	15.3	3677	21	AA550018	Stachybotrys chart
17	288	14.7	2063	21	AA550021	Curularia pallasc
18	288	14.7	2063	21	AA551316	Curularia pallasc
19	283	14.5	2905	21	AA550020	Bipolaris spicifer
20	283	14.5	2905	21	AA551315	Bipolaris spicifer
21	95.8	4.9	109519	22	AA508693	Microspora atrum
22	87.4	4.5	858	21	AA50022	Microspora atrum
23	87.4	4.5	858	21	AA51317	Microspora atrum
24	49.6	2.5	1533	22	AAH66018	C glutamicum codin
25	49.6	2.5	1614	22	AAH71220	Corynebacterium gl
26	49.6	2.5	349980	22	AAH68577	C glutamicum codin
27	47.2	2.4	523	21	AA43363	Rat secreted expre
28	47.2	2.4	523	21	AA43411	Rat secreted expre
29	44.8	2.3	10732	21	AA410594	Gene encoding a su
30	44.8	2.2	61313	23	AA559545	Protonibacterium
31	43.6	2.2	68356	22	AAK67283	Human immune/haema
32	43.6	2.2	68356	22	AAK63212	Human immune/haema
33	41.2	2.1	2012	22	AAH57385	Human skeletal mus
34	40.8	2.1	600	22	ABA59948	Human foetal liver
35	40.8	2.1	600	22	ABA28371	Probe #6837 for ge
36	40.8	2.1	600	22	AAK08219	Human brain expre
37	40.8	2.1	600	22	AAK34057	Human bone marrow
38	40.8	2.1	600	22	AAI39820	Probe #8506 used t
39	40.4	2.1	15772	22	AAK63220	Human immune/haema
40	39.4	2.0	2071	15	AAO58121	PH 2.5 acid phosph
41	39.4	2.0	2071	15	AAO58121	A. niger pH 2.5 ac
42	38.8	2.0	523	21	AA481723	N. meningitidis pa
43	38.8	2.0	20757	20	AAK20599	Polynucleotide seq
44	38.8	2.0	23457	22	AAK74866	Human immune/haema
45	38.8	2.0	31405	22	AAK67293	Human immune/haema

ALIGNMENTS

RESULT 1	
ID AAF30028	AAF30028 standard; DNA; 1958 BP.
AC AAF30028;	
XX	
DT 23-APR-2001	(first entry)
XX	
DE Stachybotrys chartarum phenol oxidase B spob gene.	
XX	
KW Phenol oxidising enzyme; phenol oxidase B; spob gene; bleach;	
KW pulp; paper; textile; detergent; ss.	
XX	
OS Stachybotrys chartarum.	
XX	
PH	
FT Key	Location/Qualifiers
FT CDS	13..1952
FT	/*tag= a
FT	1..203
FT	/note= "contains introns"
FT	/*tag= b
FT	204..255
FT	/*tag= c
FT	256..805
FT	/*tag= d
FT	806..888
FT	/*tag= e
FT	889..1628
FT	/*tag= f
FT	1629..1681
FT	/*tag= g
FT	1682..1958
FT	exon

Db	1501	gaattccgagatccttctctctcgtttccactgcgcggagagtcgagaccttaagagctgctggt	1560
QY	1561	ctcaagagatgttgcttgagctgcgtctgtctgtgagtgcttaagttaagagcccaactacgt	1620
Db	1561	ctcaagagatgttgcttgagctgcgtctgtctgtgagtgcttaagttaagagcccaactacgt	1620
QY	1621	ccattcccgtaagttctctgcgcttttaacctactggttttctaactatgcttaacatctaaca	1680
Db	1621	ccattcccgtaagttctctgcgcttttacctactggttttctaactatgcttaacatctaaca	1680
QY	1681	gtgggtgcttcaatggtttggaactgccaacctgatccagagagacacagacatgatgctg	1740
Db	1681	gtgggtgcttcaatggtttggaactgccaacctgatccagagagacacagacatgatgctg	1740
QY	1741	ctttcaatgctacatgcttctctggtgactatggtgcacaacatcacacggagttcatgaacca	1800
Db	1741	ctttcaatgctacatgcttctctggtgactatggtgcacaacatcacacggagttcatgaacca	1800
QY	1801	tgaagcctctctctgagagcccgccgccccttccctctcctctcgagaggtctcgagaaatggtctcggtg	1860
Db	1801	tgaagcctctctctgagagcccgccgccccttccctctcctctcgagaggtctcgagaaatggtctcggtg	1860
QY	1861	actctcagagagcttgccatcatcatgacacgcatccttcacagagagatggtcttaaacccctacg	1920
Db	1861	actctcagagagcttgccatcatcatgacacgcatccttcacagagagatggtcttaaacccctacg	1920
QY	1921	ccgaggtctgaatgagatgctgcgcgtcgagagaaataaacggt	1958
Db	1921	ccgaggtctgaatgagatgctgcgcgtcgagagaaataaacggt	1958

RESULT 2

AAAF82585	
ID	AAAF82585 standard; DNA; 1958 BP.
AC	AAAF82585;
DT	18-JUN-2001 (first entry)
DE	Stachybotrys chartarum phenol oxidising enzyme nucleotide sequence.
KW	Stachybotrys chartarum; phenol oxidising enzyme; detergent;
RM	paper production; pulp production; textile; food industry; bleaching; ds
OS	Stachybotrys chartarum.
XX	WO200121748-A1.
PN	29-MAR-2001.
XX	06-SEP-2000; 2000MO-EP08840.
PF	22-SEP-1999; 99EP-0203120.
XX	(UNIL.) UNILEVER NV.
PA	(UNIL.) UNILEVER PLC.
PA	(HIND-) HINDUSTAN LEVER LTD.
XX	Convents D, Doornink M, De Vries CH, Wang H;
PI	WPI: 2001-273462/28.
DR	P-PSDE; AAB81505.
PT	New detergent compositions comprising a phenol oxidizing enzyme useful
PT	in detergent or cleaning compositions, fiber treatment, processing,
PT	finishing or production, paper and pulp production, or in starch
XX	processing applications
XX	Example 5; Fig 1; 46pp; English.
XX	The present sequence encodes a Stachybotrys chartarum phenol oxidising
XX	enzyme. The invention relates to detergent compositions comprising one or

CC more surfactants and a phenol oxidising enzyme having at least 68
CC identity to the *Stachybotrys chartarum* phenol oxidising enzyme. Phenol
CC oxidising enzymes may be used in the detergent, paper, pulp, textile and
CC food industries. They are used for preventing the transfer of dyes in
CC solution from one textile to another during detergent washing, or in
CC modifying the colour associated with dyes and coloured compounds having
CC different chemical structures, such as in pulp and paper bleaching,
CC bleaching the colour of stains on fabric and in detergent and textile
CC applications.
XX
SQ Sequence 1958 BP; 394 A; 592 C; 491 G; 481 T; 0 other;

Query Match	99.98;	Score 1956.4;	DB 22;	Length 1958;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 1957; Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;

QY	1	ggctcatcaacaatgataagccaagatactgagacgctgtgctcgggaccttgctggtatc	60
Db	1	ggatccacacaaatgataagccaagctatacggagccgctggtcccgaccttgctggtatc	60
QY	61	ggcggcagctctgtctgatatccagatactcgcttgctgctgatatccagacatgaccttcggt	120
Db	61	ggcggcagctctgtctgatatccagatactcgcttgctgctgatatccagacatgaccttcggt	120
QY	121	ctcaccaagaggcagacgcaagctgaagctctccctctgacctgtgacgaagtgtcctctgcg	180
Db	121	ctcaccaagaggcagacgcaagctgaagctctccctctgacctgtgacgaagtgtcctctgcg	180
QY	181	atccctctctctgaagggcgcccaagagtagaagttacatcttatagactagcagagccaagct	240
Db	181	atccctctctctgaagggcgcccaagagtagaagttacatcttatagactagcagagccaagct	240
QY	241	tgtcaatcatgtgcagttaccgctctcccaacccaacacatgagagagacatcttgtactaga	300
Db	241	tgtcaatcatgtgcagttaccgctctcccaacccaacacatgagagagacatcttgtactaga	300
QY	301	gattgaaagatttaggcctctctccacccaattaccctgatatctgtagcgcggccaacatggt	360
Db	301	gattgaaagatttaggcctctctccacccaattaccctgatatctgtagcgcggccaacatggt	360
QY	361	tggatacagatggcatgtgtccccaagagacataccatcatctgtctccgttgtagcactgagagtg	420
Db	361	tggatacagatggcatgtgtccccaagagacataccatcatctgtctccgttgtagcactgagagtg	420
QY	421	tgtccgctctcgttgaaacagcggagagaaacactctcccaacagcgttccacttgtcaagctc	480
Db	421	tgtccgctctcgttgaaacagcggagagaaacactctcccaacagcgttccacttgtcaagctc	480
QY	481	ttctctctgagctctccctttgtatgtgttggtgctgagagaaactatacccaagccttgtagataca	540
Db	481	ttctctctgagctctccctttgtatgtgttggtgctgagagaaactatacccaagccttgtagataca	540
QY	541	ggatattactactataccccaacagcagcgtcggcccgagatgctttgtgttacctatgacatgacat	600
Db	541	ggatattactactataccccaacagcagcgtcggcccgagatgctttgtgttacctatgacatgacat	600
QY	601	gtccatacccgccgagaaacgccttacaatggtgtcagtgctggtgtctacaatgtatccaagaacc	660
Db	601	gtccatacccgccgagaaacgccttacaatggtgtcagtgctggtgtctacaatgtatccaagaacc	660
QY	661	ggctgaagatgcccctgnaacctccccaagcggtagtagcgaggtttatataccctctggtct	720
Db	661	ggctgaagatgcccctgnaacctccccaagcggtagtagcgaggtttatataccctctggtct	720
QY	721	gaactgcgaagcgatacaacgcgacagcagcactctctctccaccaatgagagagtttccag	780
Db	721	gaactgcgaagcgatacaacgcgacagcagcactctctctccaccaatgagagagtttccag	780
QY	781	cttctgggggtgacttatattcaagtggtgaagtgtggcccatgtgagatgcttccagatcttag	840
Db	781	cttctgggggtgacttatattcaagtggtgaagtgtggcccatgtgagatgcttccagatcttag	840

OY	841	aagatcgaatgaatgaatctgtgaatgccttaacacagtgctatcaagaagctgaagct	900
Db	841	aagatcgaatgaatgaatctgtgaatgccttaacacagtgctatcaagaagctgaagct	900
OY	841	aagatcgaatgaatgaatctgtgaatgccttaacacagtgctatcaagaagctgaagct	900
Db	841	aagatcgaatgaatgaatctgtgaatgccttaacacagtgctatcaagaagctgaagct	900
OY	901	tggccatgctcaacgctgcagccgagaaftacgcgtccgcgtccctcaacgctgcgtc	960
Db	901	tggccatgctcaacgctgcagccgagaaftacgcgtccgcgtccctcaacgctgcgtc	960
OY	901	tggccatgctcaacgctgcagccgagaaftacgcgtccgcgtccctcaacgctgcgtc	960
Db	901	tggccatgctcaacgctgcagccgagaaftacgcgtccgcgtccctcaacgctgcgtc	960
OY	961	tcaagctcttcctgcgtctgtatcttctgtaacctctggaagatctgaagaacgaattccctc	1020
Db	961	tcaagctcttcctgcgtctgtatcttctgtaacctctggaagatctgaagaacgaattccctc	1020
OY	1021	caggltcattgcgcgtctgaacggltggltcgtctgaagggccctgtttgaacatltgaacctgtac	1080
Db	1021	caggltcattgcgcgtctgaacggltggltcgtctgaagggccctgtttgaacatltgaacctgtac	1080
OY	1081	atcctctatggccgagcgctctggagaggttgttatcgactctccactccactctgcgtgcagctcc	1140
Db	1081	atcctctatggccgagcgctctggagaggttgttatcgactctccactccactctgcgtgcagctcc	1140
OY	1141	atcgatataccgaacacctctcctgtgtctgaagagctcgtggtttgaagcgttgatlttaaac	1200
Db	1141	atcgatataccgaacacctctcctgtgtctgaagagctcgtggtttgaagcgttgatlttaaac	1200
OY	1201	acgtgaacaggtcatalgctgaattcgtctgtatgaabaagtccttgagtcgcgccgacactctgag	1260
Db	1201	acgtgaacaggtcatalgctgaattcgtctgtatgaabaagtccttgagtcgcgccgacactctgag	1260
OY	1261	gtgcgtctccaaacccctcggaagtgcttccctctcccggaaggcggaacattggaaaccccgcaaac	1320
Db	1261	gtgcgtctccaaacccctcggaagtgcttccctctcccggaaggcggaacattggaaaccccgcaaac	1320
OY	1321	cccaactgtatgcagagactcttcacactcttcgcgcgtgtccaaatgaagacagtgaacatacga	1380
Db	1321	cccaactgtatgcagagactcttcacactcttcgcgcgtgtccaaatgaagacagtgaacatacga	1380
OY	1381	gttaacctctcggatgtgcggagaaacgctctgcctccgaatgtgcgcccgagacactgtgtag	1440
Db	1381	gttaacctctcggatgtgcggagaaacgctctgcctccgaatgtgcgcccgagacactgtgtag	1440
OY	1441	atctgcgcgaacttgcgaacaaactccaaacggttggagaccacacccgtttcaacatlaacctgtt	1500
Db	1441	atctgcgcgaacttgcgaacaaactccaaacggttggagaccacacccgtttcaacatlaacctgtt	1500
OY	1501	gaattcccgagatcctctctcgttctccactgcgcgtlgaagtcgagccttaagaagctgtgcgt	1560
Db	1501	gaattcccgagatcctctctcgttctccactgcgcgtlgaagtcgagccttaagaagctgtgcgt	1560
OY	1561	ctcaagaagatgtgcctgcgcgtcgtcgtcgtlgaaggttgtctatgtttgaagggcccaactgc	1620
Db	1561	ctcaagaagatgtgcctgcgcgtcgtcgtcgtlgaaggttgtctatgtttgaagggcccaactgc	1620
OY	1621	cccttcccgtaagttcctgcgccttttaacctgaatgttttactactcatgtcaaatctaacaa	1680
Db	1621	cccttcccgtaagttcctgcgccttttaacctgaatgttttactactcatgtcaaatctaacaa	1680
OY	1681	gttgggtctctacatgttgcgaatgcgaacaaactgataccagaaggaacagaagactgtgtgtg	1740
Db	1681	gttgggtctctacatgttgcgaatgcgaacaaactgataccagaaggaacagaagactgtgtgtg	1740
OY	1741	ccttcaatgttcaactgttctcgtgtgcataatgtcatcaacatacgaacgagttcatatgaccca	1800
Db	1741	ccttcaatgttcaactgttctcgtgtgcataatgtcatcaacatacgaacgagttcatatgaccca	1800
OY	1801	tggagcctctctcggagggcccgcccttccctcctcgtcggaagtgtcgaagaatgtgcgggtg	1860
Db	1801	tggagcctctctcggagggcccgcccttccctcctcgtcggaagtgtcgaagaatgtgcgggtg	1860
OY	1861	acttcaagcgagcttgccatatactgaacgcgattccagagagatgtctagcttcaacccctacag	1920
Db	1861	acttcaagcgagcttgccatatactgaacgcgattccagagagatgtctagcttcaacccctacag	1920
OY	1921	cccaagctcgtatgatgatgcgcgtgcaggaagtgaacgggt	1958

[illegible]

CC applications.

SQ Sequence 2095 BP; 437 A; 618 C; 510 G; 530 T; 0 other;

Query Match	99.48;	Score 1946.2;	DB 22;	Length 2055;
Best Local Similarity	99.88;	Pred. No. 0;		
Matches 1948; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

OY	3	atatacaaatgatatagaacgacgaagcatalatcggaacggtgagctctgagcccttgatgtaacgg	62
Db	67	aagatgacaatacgtatgacgccaagctatacggaaacggtggtcttcgggccccttgatgtaacgg	126
OY	63	cggacgctctgtatgatacgaagataccggtgtgctgtgatacgaagacatgcttccggct	122
Db	127	cggacgctctgtatgatacgaagataccggtgtgctgtgatacgaagacatgcttccggct	186
OY	123	caccaagaagcgagaaacgacgtatgctctcccttgagctctgtacgaagtgtcctgtccgat	182
Db	187	caccaagaagcgagaaacgacgtatgctctcccttgagctctgtacgaagtgtcctgtccgat	246
OY	183	ccctccctcgtgaagcgcccgaatgaataagtaatactctataagctgaagcagaacgcttg	242
Db	247	ccctccctcgtgaagcgcccgaatgaataagtaatactctataagctgaagcagaacgcttg	306
OY	243	ctaatactgtaagtaacgtctcccaaccccacaacactgtgaagagacatctgtactacagaa	302
Db	307	ctaatactgtaagtaacgtctcccaaccccacaacactgtgaagagacatctgtactacagaa	366
OY	303	tggagattgagccctctctccacacagatacctctgatactctgtgaagcggccaatgtgtg	362
Db	367	tggagattgagccctctctccacacagatacctctgatactctgtgaagcggccaatgtgtg	426
OY	363	gatacagatgcaatgtccccagaaacccaacatcatgtgtccctcgtgacgtatgaagtgttg	422
Db	427	gatacagatgcaatgtccccagaaacccaacatcatgtgtgtccctcgtgacgtatgaagtgttg	486
OY	423	tccgcttcgtgaacagcggagagaacaacctctcccaagaacggtccactgtgacgagctct	482
Db	487	tccgcttcgtgaacagcggagagaacaacctctcccaagaacggtccactgtgacgagctct	546
OY	483	tctctcgagctcccttctgtgtgtgtgtggtctgaaagaaactaacccaacgctgtggaataaag	542
Db	547	tctctcgagctcccttctgtgtgtgtgtggtctgaaagaaactaacccaacgctgtggaataaag	606
OY	543	attactactaaccccaacacagacgagcgtcccgccgaatgctttgttatacattatgacatgcca	602
Db	607	attactactaaccccaacacagacgagcgtcccgccgaatgctttgttatacattatgacatgcca	666
OY	603	ccaataccgcgcgaagaaacgctacatctgagtggtctacgacgtgtgtctacatacgtatccagaac	662
Db	667	ccaataccgcgcgaagaaacgctacatctgagtggtgtctacgacgtgtgtctacatacgtatccagaac	726
OY	663	ctgaagatgtcctctgaacacctccccaacgagctacgacgagtttgatatacccttggctctga	722
Db	727	ctgaagatgtcctctgaacacctccccaacgagctacgacgagtttgatatacccttggctctga	786
OY	723	cttcgcaagcgaatacgaacggaacgacatctcttctccaacaatggaagatttccaagct	782
Db	787	cttcgcaagcgaatacgaacggaacgacatctcttctccaacaatggaagatttccaagct	846
OY	783	tctgggggtgaagcttatctcaagtgtaagtttgagcccatcttgatgctctcaaatcctgaag	842
Db	847	tctgggggtgaagcttatctcaagtgtaagtttgagcccatcttgatgctctcaaatcctgaag	906
OY	843	gtatcgatgtatgaataatgtgtcatgctcttaacacagtgctataccaagaacggttcagccttg	902
Db	907	gtatcgatgtatgaataatgtgtcatgctcttaacacagtgctataccaagaacggttcagccttg	966
OY	903	gactatgctataaagtgtaacacggtcgcaagatgacggtcttcgcttctctcaacgctgcgcttc	962
Db	967	gactatgctataaagtgtaacacggtcgcaagatgacggtcttcgcttctctcaacgctgcgcttc	1026

QY	1963	agcccttcgctcgtgatacttctgtcactctctgagagatctcaagagacacagacttcccttcca	1022
Db	1027	acgcctctcttcgctcgtatctctgtcaactctcgtgagatctcaagacacacttcccttcca	1086
QY	1023	ggtcatactgcgcctgacaggtgagctcgtctggaagccctgttcaactgtcaactctgtacat	1082
Db	1087	ggctacttgcgcctgacaggtgagctcgtctggaagccctgttcaactgtcaactctgtacat	1146
QY	1083	ctctatgagccgagcgcctgaggaagttgttatcgcacttctccaactctcgtctgagccagctcat	1142
Db	1147	ctctatgagccgagcgcctgaggaagttgttatcgcacttctccaactctcgtctgagccagctcat	1206
QY	1143	cgatatctccgacacacttctctgtgtcgtgagagcgtctcgtgtgttgaagcttgattgataaac	1202
Db	1207	cgatatctccgacacacttctctgtgtcgtgagagcgtctcgtgtgttgaagcttgattgataaac	1266
QY	1203	tgaacaaagttcatgagcatctcgtctgtgtatgtagaagatcccttgaagctgcgcgcgaacacttctgaagt	1262
Db	1267	tgaacaaagttcatgagcatctcgtctgtgtatgtagaagatcccttgaagctgcgcgcgaacacttctgaagt	1326
QY	1263	gacctgcacaactctcgagatgttctccttctcccgagagcgcgcaactgtagaaccgcgcaaaccc	1322
Db	1327	gacctgcacaactctcgagatgttctccttctcccgagagcgcgcaactgtagaaccgcgcaaaccc	1386
QY	1323	caactgtaaacgaagatcttccactcttcgcgcgtcgtcctaattgtgaagaatggaataatcaacagagat	1382
Db	1387	caactgtaaacgaagatcttccactcttcgcgcgtcgtcctaattgtgaagaatggaataatcaacagagat	1446
QY	1383	taacctctccggaatgtctcgaagaacgcgtctgcctccgcgaatgtgcgcccgcgcaactgttgtagat	1442
Db	1447	taacctctccggaatgtctcgaagaacgcgtctgcctccgcgaatgtgcgcccgcgcaactgttgtagat	1506
QY	1443	cttgcgcgaacttgaagaacaaactccaacaggttgaatctaacaccctgttcaacatctcaactctgttga	1502
Db	1507	cttgcgcgaacttgaagaacaaactccaacaggttgaatctaacaccctgttcaacatctcaactctgttga	1566
QY	1503	cttcccgagctccctctctctcgtctccaacatgcgcgcgtggaagctcgaacctatgtagcgtctgtgct	1562
Db	1567	cttcccgagctccctctctcgtctccaacatgcgcgcgtggaagctcgaacctatgtagcgtctgtgct	1626
QY	1563	caaaagatgtctgtcgtgcctgcgtctcgtctcgtctgaaggtgtctatgtctgaagccccaactacgctcc	1622
Db	1627	caaaagatgtctgtcgtgcctgcgtctcgtctcgtctgaaggtgtctatgtctgaagccccaactacgctcc	1686
QY	1623	ttcccggaagttctcgcgccttctacactaactggttttcaactcatgcttaaaccttcaaacagt	1682
Db	1687	ttcccggaagttctcgcgccttctacactaactggttttcaactcatgcttaaaccttcaaacagt	1746
QY	1683	gggtgctcaatgattgtcacattgcacatgcacaacacttgatccacgagacccagacatgattgtgcgt	1742
Db	1747	gggtgctcaatgattgtcacattgcacatgcacaacacttgatccacgagacccagacatgattgtgcgt	1806
QY	1743	ttcaaatgtcaactgttctcgtctgtactgtgatagtgctatacaactaacacacgagttcatatgaccctatg	1802
Db	1807	ttcaaatgtcaactgttctcgtctgtactgtgatagtgctatacaactaacacacgagttcatatgaccctatg	1866
QY	1803	gagcctctctcgtgagagcccccgcgccttctctctctcgtgagaggttcgaggaatggtctcgtgagac	1862
Db	1867	gagcctctctcgtgagagcccccgcgccttctctctctcgtgagaggttcgaggaatggtctcgtgagac	1926
QY	1863	ttcaacgagccttgcccatcaactatgcacacgcaatctaaagagatggtctataacttcaaacctctacgac	1922
Db	1927	ttcaacgagccttgcccatcaactatgcacacgcaatctaaagagatggtctataacttcaaacctctacgac	1986
QY	1923	caagctgatgatgatgacgcgcctgtagagtaga	1953
Db	1987	caagctgatgatgatgacgcgcctgtagagtaga	2017

RESULT	4
AAF30029	
ID	AAF30029 standard; DNA; 2095 BP.
XX	



AF30029; 23-APR-2001 (first entry)

Stachybotrys chartarum phenol oxidase B spob gene.

Phenol oxidizing enzyme: phenol oxidase B; spob gene; bleach; pulp; paper; textile; detergent; ss.

Stachybotrys chartarum.

Key	Location/Qualifiers
CDS	77..2016
FT	/*tag= a
FT	/note= "contains introns"
FT	1..267
FT	/*tag= b
FT	268..319
FT	/*tag= c
FT	320..869
FT	/*tag= d
FT	870..952
FT	/*tag= e
FT	953..1692
FT	/*tag= f
FT	1693..1745
FT	/*tag= g
FT	1746..2095
FT	/*tag= h

US6168936-B1.

02-JAN-2001.

22-SEP-1999; 99US-0401476.

22-SEP-1999; 99US-0401476.

(GEMV) GENENCOR INT INC.

Wang H;

WPI; 2001-136715/14.

P-PSDB; AAB20097.

New phenol oxidizing enzyme, also useful in the detergent, paper and pulp, textile or food industries, especially in modifying the colour associated with dyes and coloured compounds, as well as in anti-dye transfer applications

Claim 5; Fig 3; 23pp; English.

The present sequence is that of the Stachybotrys chartarum MUC1 38896 spob gene encoding phenol oxidase B (see AAB20097). The gene was isolated by PCR using primers (see AAF30030-31) based on 2 peptide fragments of the isolated enzyme. The invention provides phenol oxidizing enzymes such as phenol oxidase B, or enzymes with at least 68% identity to it, nucleic acids encoding them, expression vectors, filamentous fungus and yeast host cells, and methods for the recombinant production of the phenol oxidising enzymes. The enzymes are useful for bleaching of pulp and paper, fabric stains, and in detergent and textile applications. They show optimal activity at pH 5-11, 7-10.5 or 8-10, and 20-60 or 20-40 degree C.

Sequence 2095 BP; 437 A; 618 C; 510 G; 530 T; 0 other;

	Query Match	99.4%	Score 1946.2	DB 22	Length 2095
	Best Local Similarity	99.8%	Pred. No. 0		
	Matches 1948	Conservative	0	Mismatches 3	Indels 0
0y	3 atccatcaacatgatcagccaagctatcgagacgtgtgctctgggacctgtatcgg 62				

Db	67	aacgatacaactgatacgccaagctacacggagccgctgagctctcgggcctctgctgtatctcgg	126
Qy	63	cggacagctctgatacgacgaatccgtctgtctgtatcgaacagacatgctctccgctct	122
Db	127	cggacagctctgatacgacgaatccgtctgtctgtatcgaacagacatgctctccgctct	186
Qy	123	cacccaagagcgacagacagcttgagctctccctcggcctctgtacgaagctccctctgcgat	182
Db	187	cacccaagagcgacagacagcttgagctctccctcggcctctgtacgaagctccctctgcgat	246
Qy	183	ccctctctctgaagcgccccaagtatgatactctataagcttagcagagccaactgtg	242
Db	247	ccctctctctgaagcgccccaagtatgatactctataagcttagcagagccaactgtg	306
Qy	243	ctaatacatctgagtaacgctccccaacccaactctggaagagacatcttgtatacagga	302
Db	307	ctaatacatctgagtaacgctccccaacccaactctggaagagacatcttgtatacagga	366
Qy	303	tggagatatagcccctctctccacaagaatctacacctctatcttgagccgagccaactgtg	362
Db	367	tggagatatagcccctctctccacaagaatctacacctctatcttgagccgagccaactgtg	426
Qy	363	gatacagatgacatgtccccaagagactaacatactgtctccctgtgcaactgagatgttg	422
Db	427	gatacagatgacatgtccccaagagactaacatactgtctccctgtgcaactgagatgttg	486
Qy	423	tcgcgtctctgaacagcgggagagaacacccctctccacaagcgcttccactgtgacgctct	482
Db	487	tcgcgtctctgaacagcgggagagaacacccctctccacaagcgcttccactgtgacgctct	546
Qy	483	tcctctcgagctccctctgtatgtgttggtctgagagacataccagcctctgctgatacaag	542
Db	547	tcctctcgagctccctctgtatgtgttggtctgagagacataccagcctctgctgatacaag	606
Qy	543	attactactacccaacaagcgagctgccccgaatgtttgtgtataccaatgagccaatgacatg	602
Db	607	attactactacccaacaagcgagctgccccgaatgtttgtgtataccaatgagccaatgacatg	666
Qy	603	ccatataccgcggaagaaacgacctatactgtgtacaagctgtgtgtctatactgataccagagacgg	662
Db	667	ccatataccgcggaagaaacgacctatactgtgtacaagctgtgtgtgtctatactgataccagagacgg	726
Qy	663	ctgagagatgcccctgtgaacccctcccaagcgctcaagcgagatgttatatccctctgtgtctga	722
Db	727	ctgagagatgcccctgtgaacccctcccaagcgctcaagcgagatgttatatccctctgtgtctga	786
Qy	723	ctgccaagaacgataaacaagcagacgcactctctctccacaacaaatgagagagttctcaact	782
Db	787	ctgccaagaacgataaacaagcagacgcactctctctccacaacaaatgagagagttctcaact	846
Qy	783	tcctgggtgtagcgttatctcaagtggttaagttagagccatctgagatgctctcaagatccctagaa	842
Db	847	tcctgggtgtagcgttatctcaagtggttaagttagagccatctgagatgctctcaagatccctagaa	906
Qy	843	gtatcagatgataatgttgacatgtctctcaacaagtgctataccaagaaggtgtcaacgttg	902
Db	907	gtatcagatgataatgttgacatgtctctcaacaagtgctataccaagaaggtgtcaacgttg	966
Qy	903	gactatgtctacaagtgtagcgcgcgccaagtataccgctctccgctctctctcaacgctgtccgctctc	962
Db	967	gactatgtctacaagtgtagcgcgcgccaagtataccgctctccgctctctctcaacgctgtccgctctc	1026
Qy	963	acgctctcttgctctgtatattgtcttaacctctgtagatctcaagagaccaagatctcccttcca	1022
Db	1027	acgctctcttgctctgtatattgtcttaacctctgtagatctcaagagaccaagatctcccttcca	1086
Qy	1023	ggtcaattccgctgaacggtgtctgcgtcttgtagggccctctgtgacaactgtgaacactgttaaat	1082
Db	1087	ggtcaattccgctgaacggtgtgtctgcgtcttgtagggccctctgtgacaactgtgaacactgttaaat	1146
Qy	1083	ctctatacgccagagcgtctgggaagtggttatacgactctccacactctgcgtgtgccaagctcat	1142
Db	1147	ctctatacgccagagcgtctgggaagtggttatacgactctccacactctgcgtgtgccaagctcat	1206

OY	779	agttctggaggtgagcgtctctcaacgagtgtaagtgtgagccattggaggtcttaagatctc	838
Db	771	tcattctgggggtgtgtatcaatcaagc-----	796
OY	839	agaagatcgaatgtatgaanaatvtgtgatctctaaacagtgctatcacagaacgttcagc	898
Db	797	-----gaacgttaac	807
OY	899	cttggcctatgctcaagvtgcacgcgcgcgaagacgcgtcttcgcttctctcaacgctgcgc	958
Db	808	ccgtgaccttcaagaagcgttgagccctccgcgaataatgatattccgcttccctcgatgtgcgag	867
OY	959	tcctacgcctcttcgcctctgtatctgtctaaaccttcagagattccaagacacagattccct	1018
Db	868	ttctcgcctcttcgcgccttactctgtctgcgtatcagatgtctatcgcacatcgtctgctct	927
OY	1019	tccaggtcatctgcgcctgcagcgtgtgtctctgtcgtatgagggccctgtgtacacatgcactgt	1078
Db	928	tcaaggttatgtccctccgcatctctgtctctctttgaacacaccttcgcgataccagctgtgcgt	987
OY	1079	acactctcatgcccgcagcgcgtgtggaggtgtgtatcgcactctccacacttcctgcgcgcgt	1138
Db	968	acatttccatgcccgcagcgtgtacgaaggtgtgtgtttgacttccctccgacatgctgtgcaga	1047
OY	1139	ccatcgaatcgcgcgaacctctctcgtgtg---ctgacggtctcgtgtgtgtgagccatgattgt	1195
Db	1048	ctattgacatctccgcgaacctgtggcgcgtgacgatctgtgcgcacatcgacagatcacgcattgt	1107
OY	1196	ataacatcttgacaaaggtcatgtcgaattctgtctgtgtatgagaatctcttgatgcgcgcgaacct	1255
Db	1108	acaaacacgcgaagaaggtcatgtcgtctcgtgtgtatgacgaacgcacaaactcaagccagataacct	1167
OY	1256	ctgagcgtgcctgcgcgaacctccgcgaatgtctcttcccccgcagggcgcgcgaactgtggaccgcg	1315
Db	1168	cagttgtctcgtcaactctcgtatgtctcccttccctccctccacacaaacacccccc	1227
OY	1316	caaaaccccatcgtatgacagacttcaactcttcgcgcgcgcgtctatgtacatgacagtgacataca	1375
Db	1228	gaca-----gtctcgcctttgtgtctgcacgcggtctctacatcgtgagctatta	1269
OY	1376	acggagatctactctctcgtatgtcgaagaaacgcgtctgcctccgaatgtgtgcgcgcgacaaact	1435
Db	1270	atggtgtgtcgtcttgcgtatgcatgtccaacacgcgtcgtctgtcgaacgttaccgcttgtactgt	1329
OY	1436	cttgagatcctgcgcgaactgtgaagaacactccaacgcggtgtgacatcaaccgttcaacatcaacc	1495
Db	1330	tcgagcgtctggagatcatalcaaacgcgcgtgaacggtgttgagacgccttatcaatcatc	1389
OY	1496	tcgctgtgactccgaagtcctctcttcgcttc-----cacgtccgcgttgagtgagacgct	1546
Db	1390	ttgtgcgacttccaaggtgatattctcgttaacttccgcgaacaaacgcgcgcgaacgtcatgtccat	1449
OY	1547	atgagcgtctcgtgtctccaagaagatgtatctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1606
Db	1450	acagc---tccggtctcaagaagcgtgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt	1506
OY	1607	agggccactcagctctcctcttccogtaagcttccgcttctactaaactggttttcaatcat	1666
Db	1507	aggtctcactaagcgcctcttccc-----	1528
OY	1667	gctaaatctcaacagtcgtgtctatacgttctgacgtcccaacccgtatctcacacgagagcca	1726
Db	1529	-----cgtgtataacatgttctcatatgtccgaacaaattgtatctcaacgagatca	1573
OY	1727	cgacatgacgtgctgtcttccaatgtcaactgttctcgcgtgacatatgtcgtctacaacaccca	1786
Db	1574	cgaatgatgtgcgtcctttaaaccacgcgtctctgcacagattatgtactaataatgcccactgt	1633
OY	1787	gttcatgtgacccaatgagagcctctcttgagagggccgcgcgcgcctctccctccgagagagctga	1846
Db	1634	ttctcgttgacccctatggaagagcctcttggagacgcgtcgtccctctatgtacacgtgcgcgtttcca	1693

Qy	1897	gaagtcggtgggagctacagcaggttccatcctacagccgcatcagagagtgctag	1906
Db	1694	ggtcccaagtagggcgagcttcaagcttcaagctgttactcagagctattccagactaagcttga	1753
Qy	1907	cttcaacacctagcccaagctgatatga	1933
Db	1754	atacagagaccttaccgacagcttgaagca	1780

RESULT 6

ID	AAZ61243	standard; DNA; 2110 BP.
----	----------	-------------------------

AC AAZ612

DT 30-MAY-2000 (first entry)

DE DNA encoding a phenol oxidising enzyme.

KW Phenol oxidizing enzyme; fungus; redox reaction; detergent; bleaching;

KW coloured compound; porphyrin; tannin; polyphenol; carotenoid;

KW anthocyanin; Maillard reaction product; ss.

OS Acremonium murorum.

FH	Key	Location/Qualifiers
125	1043	

FJ
DUE

XX
FY E000000E3A0 97

XX 03-FEB-2000
DD

XX 12-THT-1000, 6060-EP04033
PE

XX 21-THI-1008, 00FB-0303454
DP

XX
XX
PA (UNIT) UNIT EVER NY

PA (UNIL) UNILEVER PLC.
PA (HIND-) HINDUSTAN LEVER LTD

XX	Convents D	Goulka PT	Van Der Heijden M	Szwachhoff T	Vorjins CM
PT					

XX
DB
WPT: 2000-195101/17

DR P-PSDB; AAY69204.
XX

PT phenol-oxidizing enzyme from *Acremonium* for bleaching stains on fabrics -

XX
PS
Example 2: Page 37-40: English

Pharmacokinetic and pharmacodynamic studies

CC *tungus Acremonium murorum*. The enzyme has the CBS accession number
CC 157 72 The enzyme catalyses redox reactions and is specific for

is specifically used in detergents for bleaching strains on fabrics

CC but also for bleaching pulp and paper and for decolourisation of
CC plant-derived food products. The enzyme has a pH optimum in the

alkaline to neutral range and can bleach a wide variety of coloured compounds e.g. porphyrins, tannins, polyphenols, carotenoids.

CC anthocyanins and Mallard reaction products.
XY

Sequence 2110 BP; 400 A; 732 C; 556 G; 422 T; 0 other;

Query Match	26.48;	Score 517.4;	DB 21;	Length 2110;
-------------	--------	--------------	--------	--------------

Best Local Similarity 61.4%; Pred. No. 4e-132;

Matches 1041; Conservative 0; Mismatches 476; Indels 178; Gaps 6;

257 accgtccccaacccaactgagagagacatctgttactacagatgagattagccc 316

[illegible][illegible]

RESULT	7
AAZ27601	
ID	AAZ27601 standard; DNA; 1791 BP.
XX	
AC	AAZ27601;
XX	
DT	16-DEC-1999 (first entry)
XX	
DE	Stachybotrys phenol oxidase coding sequence.
XX	
KW	Phenol oxidase; enzyme; coloured compound; dye transfer prevention;
XX	fabric washing; stain bleaching; anti-dye transfer; detergent; ss.
OS	Stachybotrys chartarum.
XX	
PN	W09949020-A2.
XX	
PD	30-SEP-1999.
XX	
PF	23-MAR-1999; 99WO-US06327.
XX	
PR	24-MAR-1998; 98US-0046969.
XX	22-DEC-1998; 98US-0218702.
PR	22-MAR-1999; 99US-0273957.
XX	
PA	(GEMV) GENENCOR INT INC.
PI	Amory A, Wang H, Dhase P, Lambrechts-Rongvaux A, Wang C;
XX	
WI	WI1; 1999-551088/50.
DR	P-PSDB; AA939992.

[illegible]

Db 1169 gtgacgtctcttcc-----tctcaagaaggccgcgcgaagc 1213
 QY 1337 cttcaacttcgcgcgtgtaattgagacagtgagacaatcaacgagttactctcgatg 1396
 Db 1214 acttcaagtttgaacgcagacagcagacacattacatcaagatgttgcttgcgatg 1273
 QY 1397 tggagaacgcgtctcgcgaattgcccgcgcgaactgttgatctgacgacttga 1456
 Db 1274 tcaatgacgtgtctcgcgaaggccgcgcgacgttgacacgttgatctgagatcgaga 1333
 QY 1457 acaactcaacgagttgtaactacacttctacattacactgttgatctgcgaatctt 1516
 Db 1334 actctctgc 1333
 QY 1517 ctgcgtcactgc 1573
 Db 1394 agcgaaactgt 1453
 QY 1574 tctgt 1633
 Db 1454 tctgt 1508
 QY 1634 tctgcgccttacttaactggttttcaatcattgataacttaacatgagtgatcat 1693
 Db 1509 -----tgagacttcaat 1520
 QY 1694 gtgacatgcacacactgtatccacgagacacacacacacacacacacacacacacac 1753
 Db 1521 gtgacatgcacacactgtatccacgagacacacacacacacacacacacacacacac 1580
 QY 1754 tgtctgt 1813
 Db 1581 cgcacatgag 1640
 QY 1814 gaggccgc 1873
 Db 1641 gc 1700
 QY 1874 tgcacatgcac 1933
 Db 1701 gtccatcactgc 1760

RESULT 10
 AA25727
 ID AA25727 standard; cDNA; 1791 BP.
 XX
 AC AA25727;
 DT 05-JAN-2000 (first entry)
 XX
 DE Stachybotrys chartarum phenol oxidising enzyme encoding cDNA.
 XX
 KW Stachybotrys chartarum; phenol oxidising enzyme; colour; dye;
 KW detergent; anti-dye transfer; stain removal; bleaching; ss.
 XX
 OS Stachybotrys chartarum.
 PN W09949010-A2.
 XX
 PD 30-SEP-1999.
 XX
 PF 23-MAR-1999; 99WO-EP02042.
 XX
 PR 24-MAR-1998; 98US-0046969.
 PR 22-DEC-1998; 98US-0218702.
 XX
 PA (UNITL) UNILEVER NV.
 PA (UNITL) UNILEVER PLC.
 PI Convents D, Amory A, Wang H, Dhaese P, Lambrechts-Rongvaux A;
 XX Wang C;
 XX

DR WP1; 1999-601211/51.
 DR P-PSDB: AAY45222.
 XX
 PT Detergent composition containing phenol oxidase from Stachybotrys, used
 PT to bleach stains and prevent dye transfer
 XX
 PS Example 15; Fig 5; 56pp; English.
 XX
 CC The present invention describes a detergent composition containing a
 CC purified phenol oxidising enzyme derived from Stachybotrys. The present
 CC sequence encodes Stachybotrys chartarum phenol oxidising enzyme. The
 CC enzyme can be used to modify the colour of dyes and other coloured
 CC compounds (e.g. for use in pulp and paper bleaching also for removing
 CC stains, e.g. food, tea, blood etc., from fabrics) and for preventing dye
 CC transfer during fabric washing.
 XX
 SQ Sequence 1791 BP; 380 A; 554 C; 448 G; 409 T; 0 other;

Query Match 23.4%; Score 458.8; DB 20; Length 1791;
 Best Local Similarity 59.2%; Pred. No. 5.2e-116;
 Matches 995; Conservative 0; Mismatches 522; Indels 163; Gaps 6;

QY 257 acggtccccaacccaacactgtgagagacatcttctgtaactgagatgagatgagccc 316
 Db 241 atcataccaacccctgtcaaccgcgaagacattgttactatgatacgaatcaagca 300
 QY 317 tttccccaacgacatccctgtatctgtgagccgcgaacatggttgatagatgagcgt 376
 Db 301 tttaagaagaatttaacacactgtgcctgtgcacctcctgcgcgtgataagatgagc 360
 QY 377 tccccaacacatcaatcattcctcctgtgtgacactgagatgtgtgtcgcctgtgaac 436
 Db 361 agccctgtctacttcaatgttccacgcaggaacacagactgtagttagttcataaac 420
 QY 437 agcgagagaaacactctcccaacagcgtccacttgcacagcgtcttctctcgagccc 496
 Db 421 a-----atgcacacgtgagagacatcgtccatctgacagcgtcccatcgctgcct 474
 QY 497 tttagt 556
 Db 475 ttcgatgt 534
 QY 557 aacagagagcgtgc 616
 Db 535 aactacaatccgc 594
 QY 617 aacgcctacatgggttcagagcgtgtgtctacatgtatccagacccgcgtgagatgcct 676
 Db 595 aatgcctacttgttcagagcgtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 654
 QY 677 aactcccccgc 736
 Db 655 ggtctccctgt 714
 QY 737 aacgcagagcgaactctctctccacacatgagaggtttccagctctcgtgagcgtt 796
 Db 715 aacgcgatgttaccctcgt 774
 QY 797 attcaagtggtgaattgagcccatltagatgtctcagatcctagaaatagatgata 856
 Db 775 atccatgtc----- 783
 QY 857 aatgtgtatgtcttaacacagtgctatcaacagacggttcagccttgctcatgtcaag 916
 Db 784 -----aacgcagacacatgagccttctcctaag 811
 QY 917 tgcagccgcgaagtacacgcttccctcctcaacgcgtgcgcgtcctctcgcctc 976
 Db 812 tccagcccggaagtacacgcttccctcctcaacgcgtgcgcgtcctcgcctc 871
 QY 977 tgtatctgtactcctgtgagatcagagacagactcctcctcaggtcaattgtgcgtg 1036

```
Db 872 tctacctgcgcgaagaccagctctcccaagctcaacgtcttcccaagtcattgctctg 931
QY 1037 acggtggtcgtcttgagggccctgttgacactgacactgttaactctatggtcgagc 1096
Db 932 atgctgcctccctcaagccccctctagaaccttaaccttaacctgtgtgtgcgagc 991
QY 1097 gctgaggggtgtgtatgatctctccacctctgctgacgagtcacatgcagacac 1156
Db 992 gttacagagatcattatgatcttcaacaaacttctgtgcagactcttgacctgagcag 1051
QY 1157 ttccctggtctgaacggtctggtgtgtgagccttgagcttgatgaactgacaagtgatgc 1216
Db 1052 ttgctgagacaacagatgctgacagagatgagtaagctgcagactctcaggtgtgtgc 1111
QY 1217 gattcgtcgttgaagaagtctctgagtcgcccagactcttgaggtgtgctgcgcaactcc 1276
Db 1112 gcttcgctgcgcagctgacgtgtgtgag---gacaacagcaaggtccctccactcc 1168
QY 1277 gagatgtctcttcccccagggcgcaactgaggaccgcgaanaacccactgtatgcaga 1336
Db 1169 gtgacgtctcttcc-----tctcaacaaggagcccgccgacagaagc 1213
QY 1337 cttcaccttcgcgcgtgtcaatgagacagtggacaatcaacggagtaacctctcgatg 1396
Db 1214 actcaagtttgaacgagacagacgacactaacctgatacagatgtgtgttgcgagtg 1273
QY 1397 tggagaaacgtctgctgcgaatgtgcccgcgacactgttgatgttgagatgtgcgagttaga 1456
Db 1274 tcaatgagcgtctcctgcgcacagccgagctgcgacactgttgagcttggagagctgcaga 1333
QY 1457 acaactccaacgggttgagactaacctgttacaattcaactcgtgtgacagagctctt 1516
Db 1334 actcctctgagggctgagacaccccgctcccaattcaactgtgtgattcaagctccta 1393
QY 1517 ctgtgtccaactgcgcgcg---agtgcagccttatgaggtctgctggtctcaagatgttg 1573
Db 1394 agcgaactgtgtgtcgtgcgaagtgatgccttaagtgatgtctgtgtcttaagatgtgc 1453
QY 1574 tctgctgctgcgtcgtgcgaggtgtgtctatgttgagggcccaactacgtctctccgtaag 1633
Db 1454 tctggttgggcaggggtgagagacctgacactcgagggcccaactcaacctgtgagc----- 1508
QY 1634 ttctgccttttaactaaactggttttaactcaatgatacaactcaagaagtgtgtctaacat 1693
Db 1509 -----tgtagcttacct 1520
QY 1694 gttgcactgcaacaactgtatccagagagacacagacatgattgctcttcaatgcac 1753
Db 1521 gtggcaactgtccacaacctcaatcaagagabaaagacatgattgtgtatcaacgctcac 1580
QY 1754 tgttctcgtgtactatgtgctacaactaacacaggttcatgtgaccccaatggagcctctctg 1813
Db 1581 cgcatacgtgagagaaagatatcttcaagagagacttcgagagaccatgtaaccccaagtgc 1640
QY 1814 gaagcccgccctctctctcgtgcgagaggttcgagaatgctgcgggtgacttcaagagact 1873
Db 1641 gcgcgcgttctcttaacaaccccaacgacttccatgtcgcgctggaacttctccgcga 1700
QY 1874 tgcctacactgacacgcatcattagagagatgctgacttcaaaccccaatgcagagtgta 1933
Db 1701 gtccatcaatgcgcagagtcgagagctgcgcgagcagagacgcgtlaaacgcgctcgatga 1760
```

RESULT 11

AA25735 standard; DNA: 2067 BP.

```
XX AA25735;
XX 05-JAN-2000 (first entry)
DT Stachybotrys chartarum phenol oxidising enzyme PCR fragment.
DE
XX
```

```
KM Stachybotrys chartarum: phenol oxidising enzyme; colour: dye;
KW detergent; anti-dye transfer; stain removal; bleaching; PCR primer; ss.
XX Synthetic.
OS Stachybotrys chartarum.
XX MO9949010-A2.
XX 30-SEP-1999.
XX 23-MAR-1999; 99MO-EP02042.
XX 24-MAR-1998; 98US-0046969.
XX 22-DEC-1998; 98US-0218702.
XX (UNITL ) UNILEVER NV.
PA (UNITL ) UNILEVER PLC.
PI Convents D, Amory A, Wang H, Dhasee P, Lambrechts-Rongvaux A;
PI Wang C;
DR WPI; 1999-601211/51.
XX PT Detergent composition containing phenol oxidase from Stachybotrys, used
XX to bleach stains and prevent dye transfer
XX Example 17; Fig 9; 56pp; English.
XX The present invention describes a detergent composition containing a
XX purified phenol oxidising enzyme derived from Stachybotrys. The present
XX sequence represents a PCR fragment of Stachybotrys chartarum phenol
XX oxidising enzyme. The enzyme can be used to modify the colour of dyes
XX and other coloured compounds (e.g. for use in pulp and paper bleaching
XX also for removing stains; e.g. food, tea, blood etc., from fabrics) and
XX for preventing dye transfer during fabric washing.
SQ Sequence 2067 BP; 445 A; 621 C; 500 G; 501 T; 0 other;
Query Match 15.3%; Score 300.4; DB 20; Length 2067;
Best Local Similarity 55.2%; Pred. No. 2.4e-72;
Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;
QY 316 ctctcccaacatctacacctgatactgagccggccaacatgttgatagatggat 375
Db 457 ctgaccttccagattaccccaacctgcgcctgcacactctgctgcagatggat 516
QY 376 gtcccaagacatcatctatctctctgctgacatgagagtggtccgtctgtgaa 435
Db 517 gaagccctgctcaacttcaatgttcccaagagaaacagagactgttagttacatcaa 576
QY 436 cagcgagagaacacctctcccaacagcgctcaacttgcagagctcttctcgaagctc 495
Db 577 ca-----atgcacagtgtagaagactcgttccatctgcagagctcccaatcggtgcc 630
QY 496 ctgtatgtgtgtgctgtagaacataccacagcctgagagtaaaagatatactaaccc 555
Db 631 ttctgatgtgtgtgtagaagatgtgaccttccctcgtggagtagaagattactacttcc 690
QY 556 caacagcgaggtcgccgcagatgtgtgtgtaacatgacatgc----- 597
Db 691 caactacaatccgcgcgctctctgtgtacatgacacagcttcaatgaagatagcta 750
QY 598 -----catgtccatcacccgcgag 616
Db 751 cgagccttattcttctgtactacttggctaaccaactcttctgtgactgtcgag 810
QY 617 aacgcctacatggtcagagctgtgtgtatcatgatccagagcccggtgagatgcctg 676
Db 811 aatgcactacttgttcagagctgtgcgtacatacatcaacagcagagctgtagtctctc 870
QY 677 aacctccacgcgctacagcgagttgatatacccttggttctgactgcgaagatgac 736
```

```

Db      871 gctcttcctagtgatgcatgagatcagatataccctcttgatccttgagccaaagtactat 930
QY      737 aacgcagacgcacactctctctccaccacatgagagttccacgtcttggtgagctt 796
      931 aacgcagatggtaccctcgctctgacccgagggagagacagacactggtggagagatgct 990
QY      797 attcaagtgttaagtgtgagcccatgtgacatgtctcagatcctagaagatcatgtatga 856
      991 atccatgtc----- 999
QY      857 aattgagcatgctcttaaccagttgatacacagaaagtcagccttgagcctatgctaaag 916
      1000 -----aacgagcagacatgagccttccctcctc 1027
QY      917 tgcagccgcgcaagtaacgctctccgctctccactcaacgctgcgctcagcctcttcgctc 976
      1028 tccagcccccgaagtaacgcttccgattcctcaacgctgcgctcgtctcgtcttgctc 1087
QY      977 tcatcttgtaacctctgaaagattcagagacagacacttccctccaaagtcatgtgcgctg 1036
      1088 tctaccctgcagagacagactctcccaagctcagaattcctctccaaagtcatgtgcctc 1147
QY      1037 acggtgtctgctgtgagggccctggtgacactgacactctgtatcatctctcatgcccagc 1096
      1148 atgtgtgtcctcctcaagcccccgttcagacctcttaaccttaacctgtctgtgcgagc 1207
QY      1097 gctggagagttgta----- 1111
      1208 gttacagagatcattatgtatgctcctccctccctcaacgaatgagtcagaacctctaaagc 1267
QY      1112 -----tcgactctccactctgctgccaagtcacatcagatalccgaacctctc 1162
      1268 taacctgtgagactctcaacacttctgccaagactcttgacctgccaagctgtcgt 1327
QY      1163 gtgtcgaagctctcggtgtgtgagcgtgagtttataaactctgacaaggtcatgagatcgt 1222
      1328 agaccacagatgtctgcgagagatgtatgactcgcacactctcgaagtgatgctgc 1387
QY      1223 tctgttagaagctcgtgagctgcgccacactctctgagtggtgcgccaactccgaagatg 1282
      1388 tctgacagctctgacactggtgag-----gacaaagccaagtcctccctcaactccgtagag 1444
QY      1283 ttccttccccgagggcgcaactgtggacccgcgaacccccactgtatgacgaacttca 1342
      1445 ttccttccct-----cctcaagaaggcccgccgacagaacttca 1489
QY      1343 ccttgcgcgtgtctatagtcacgtgtgacaatcaacgaggttaacctctcgatgtcgaaga 1402
      1490 agtttgaacgcagacagagacactaactgtacaagatgttgccttgcgcatgtcatag 1549
QY      1403 accgtctgcgcgcaatgtgccccgcgacacactgttgagatctgagactgagaaacaact 1462
      1550 agcgttctcgcgcaagcccgagactcggacccgttggagctcgtggagctcgaagaacctc 1609
QY      1463 ccaacggttggactcaacctgttacaattcaactcgttgaactcgcagctcttctcgt 1522
      1610 ctgagagctgagacacaccgcctcacattcaactctgttgaacttcaagatcctaaagaa 1669
QY      1523 ccaactgcgcgtg-----agtcgaacctatagagctgctgttccaaagatgtgtctgc 1576
      1670 ctgtgtgtctgtgcaagtgatccctacagatctgtgtcttaagatgagtcgtcgtgt 1729
QY      1580 tggctcgtctgagagttgtctatgttagagccacactacgctcttccctccgtaagtctcg 1639
      1730 tgggacaggggtgagacacctgtacatcagagcccaactcaacctgtgac----- 1778
QY      1640 cctttaactaactggttttcaactcatgtacaatcacaagtgtgtctacatgttgcga 1699
      1779 -----tgagacttaaatgtgcga 1796
QY      1700 ctgcacacactgatccagagagccagacatgagctgcttcaatgttaactgtctc 1759
      1797 ctgtcaaacactcatltaacgagataacgacatgtatgtctatltcaacgtccacgcgcat 1856

```

```

QY      1760 cgttgactatgctacactacacacagatcattgaccccatgagcctctcgaagcc 1819
      1857 gggaggaagaaggatattcttaagagagacttcgaggaacctatgaacccaagtgggcgcc 1916
QY      1820 cgcgcctctcctctcctcgagagatcgaagaatggtcgtggtgaactcaacgagcttgcac 1879
      1917 cgttcttaacacgcagacactcattcatgtctgcgttggaacttctcgcgagctcact 1976
QY      1880 cactgcagcatcagagagatggttagcttcaaccttaagcccaagctgatatga 1933
      1977 cactgcgcagtgatgaagagctgagccgagcagagagcgtataacccgcgtcatga 2030

RESULT 12
AAZ27609
ID  AAZ27609 standard; DNA: 2067 BP.
XX
XX  AAZ27609;
AC  AAZ27609;
XX
XX  16-DEC-1999 (first entry)
DT
XX
DE  Stachybotrys phenol oxidase coding sequence.
XX
XX  Phenol oxidase; enzyme; coloured compound; dye transfer prevention;
KW  fabric washing; stain bleaching; anti-dye transfer; detergent; ss.
XX
XX  Stachybotrys chartarum.
OS
XX
XX  WO949020-A2.
PN
XX
XX  30-SEP-1999.
PD
XX
XX  23-MAR-1999; 99MO-US06327.
PF
XX
XX  24-MAR-1998; 98US-0046969.
PR  22-DEC-1998; 98US-0218702.
PR  22-MAR-1999; 99US-0273957.
XX
XX  (GENEV ) GENENCOR INT INC.
PA
XX
XX  Amory A, Wang H, Dhase P, Lambrechts-Rongvaux A, Wang C;
PI
XX  WPI; 1999-591088/50.
DR
XX
XX  Novel enzyme for modifying coloured compounds used to prevent
PT  dye-transfer .
PS
XX
XX  Example 17; Fig 9; 64pp; English.
XX
XX  This sequence encodes the Stachybotrys chartarum phenol oxidase enzyme
CC  of the invention. The invention is used to modify a coloured compound and
CC  prevent dye transfer during fabric washing, or for stain bleaching or
CC  anti-dye transfer. It is useful in the detergent, paper and pulp, textile
CC  and food industries.
XX
XX
SO  Sequence 2067 BP; 445 A; 621 C; 500 G; 501 T; 0 other;

Query Match      15.3%; Score 300.4; DB 20; Length 2067;
Best Local Similarity 55.2%; Pred. No. 2,4e-72;
Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;

```

```

QY      316 ctctccacacacatcaacctgtatcgtgagccgcgcaacatgttgataagatgcat 375
      457 ctgaccttcaagatthacacacacacttgccctgcctgcacactctcgtgcatgagatgcat 516
QY      376 gtcccagacactacatcatcgtctcgttggaactgaaagtgttcgcctctgaa 435
      517 gagccctggtccacttcaatgttcccgagagaaacagacagactgtatgtatcatcaa 576
QY      436 cagcggagagaacacactctcccaacagcgtccacacttgacagctcttctcagagctcc 495
      517 ctgtcaaacactcatltaacgagataacgacatgtatgtctatltcaacgtccacgcgcat 1856

```



```

Db 577 ca-----atgccacgcgtggagaactgcgtccatctgcagcgtcccatcgctgcc 630
QY 496 ctttgatgcttgagctgaaggacaactaccagcgtgagtgataagattactactccc 555
Db 631 ttccagatgcttgagctgaaggatgtgacctccctgcgcgagtaacaagattactactcc 690
QY 556 caacaggcagcgtgcgcacatgtcttgtaaccatgacaatgc-----597
Db 691 caactacaataccgcgcgcctctctgtgtacatgacaacgccttcaatgaagtacta 750
QY 598 -----catgtccatcaacgcgcag 616
Db 751 cgaagccttaactctctgtgtacatctgtgtacacaactcctctgtgtagatgcgtga 810
QY 617 aagccctacaatggtgacgctgtgtgtatcatgatcagaacccgcgtgtgagatgcctg 676
Db 811 aagccctacttgtgcagcgtgtgcgcctacaatacacaagcagcgtgtgagatgcctc 870
QY 677 aacctcccaagcgtgctacgagcttgatatacccttggcttctgactgccaagcgaat 736
Db 871 ggtcttctagtgctatgagcaggttcgataccctctgataccctgcgcgcgaagtact 930
QY 737 aagcgaacagcgcctctctctccacaatgagaggtttccacgctctctggtgagctc 796
Db 931 aacgcgaatgagcagcgtgtgcgcacgaggtgaggaacagacccgtgtggtgagatgc 990
QY 797 atcaaatggttaagtgtgagccatgtgatagtctcagatcctagaagtatcgatgtga 856
Db 991 atccatgtc-----999
QY 857 aattgtgcatgcttaacagatgctatacaagaagctgacgcttgcctatgctcaag 916
Db 1000 -----aagcagacagcagcagcttcccttccaaag 1027
QY 917 tgcagccgcgaagtacagcttccgcttcccaacgctgcgtgtcagcgtctctgtgc 976
Db 1028 tccagccgcgaagtacagcttccgcttcccaacgctgcgtgtcagcgtctctgtgc 1087
QY 977 tgaatctgttaacctctgagagatcagaagacagacttccctccaggtacatgtgcgtg 1036
Db 1088 tcaactctgtcaagagacagcttcccaacgcttcagaatctcttcccaagtaattgtcc 1147
QY 1037 aaggtgtctgcttgagagccctgtgtgacacatgaacctgtatcatctctatgcccagc 1096
Db 1148 atgctgtctcctctcaagcccccgttcaagccttaaccttaacctgtgtgtgcagc 1207
QY 1097 gctgcgagaggtgtta-----1111
Db 1208 gttacagatcaltatgtgatagcctccctccctccacgaatgagtaagaacttaagac 1267
QY 1112 -----tcgaacttccaccttgcgtgtgcagttccatcgatataccgaaccttctg 1162
Db 1268 taacactgtagacttcaaccaacttgtgcgcgaacttctgacctgtgcgaagctgtg 1327
QY 1163 gtgtcgaagcttcgtgtgtgtgagcgtgtgtataaactgaagaaggtcatgcatgcgt 1222
Db 1328 agaccacaagatgtcgcgagagatgtagatcgtcgaactctcagaggtgtatgcgtc 1387
QY 1223 tctgttagatgagctcgtgagctgcgcgcacactctgtgagtgagcgaacctccgagatg 1282
Db 1388 tctgtaagctctgtgacactgtgtgag-----gacacagcagcaggtccctccactctccgtgagc 1444
QY 1283 ttcccttcccgagggcggaactgtgagaccgcgaaccccaactgtatgacagagacttca 1342
Db 1445 ttcccttccct-----cctcaagaagagcccgccgcgcgaacaagacttca 1489
QY 1343 ccttgcgcgtgtgtaatgagcagatgacaatcaacgagatgaacctctgcgagtgtcaga 1402
Db 1490 agtttgagcagcagcaagcagacactgaactgaacgagtgtgtgtgtgcgagtcatg 1549
QY 1403 accgtctgcgcgcgaatggtgccccgcgacactgttgagatctgagcagcttgagaacact 1462
Db 1550 agcgtctgcgcgcgaacccgagctgcgacgcttgagcttgagagctcagagaactcct 1609

```

```

QY 1463 ccaacggttgactcaaccctgttcaactcaactcgtgtgactccgactccttctcgtt 1522
Db 1610 ctgagagcgtggaacaccccgctccacactcaacttgtgtgacttcaagatctcaagcgaa 1669
QY 1523 ccaatgcgcgttgg---agtcgaacctatgagcgtgtgtgttccaaagatgtgtcgtgc 1579
Db 1670 ctgtgtgtcgttgccaggtcatgtccctacgagctctgtgtgttcttaagatgtcgtcgtg 1729
QY 1580 tgcgtcgtcgttgaggtgtgtatgtgttgaggtccactacgctcctctccgttaagttcgt 1639
Db 1730 tggcgaagggttgagaccctgacatcgagcccaactcaacccttgcagc-----1778
QY 1640 cctttacctaactggtttcaactcatgatactcaactcaagaatggtgtctacatgttga 1699
Db 1779 -----tggagcttaactgtggca 1796
QY 1700 ctgcacaacctgataccaagagacacgaacatgagctgtcttcaatgtcaactgttct 1759
Db 1797 ctgtcacaaacctaatctacgagagtaaacgacatgagtgtctgtatccaagctcacgcgcat 1856
QY 1760 cgtgtgactatggttacaactacaacgcaggtcattgaccccatgagcgtctctggagcgc 1819
Db 1857 ggaaggaagagatatacttcaagagagacttcgagacccaatgaaaccaagtggcggc 1916
QY 1820 ccgcccctctcctcctcgcgagagcttcgaaatgtgcgtcggtgacttcagcgaacttgcat 1879
Db 1917 cgttcttcaacacgcgaacgaacttccatgtcgtcgtggaacttctccgcgcgagttccat 1976
QY 1880 cactgacgcgcatccaagagatgtgacttcaaacctcaacgcccgaagctbatga 1933
Db 1977 cactgcgcgcgagtgacagagctgtgcgagcagagccgttacaaccgcgtcgtatga 2030

RESULT 13
AAZ25728
ID AAZ25728 standard: DNA: 3676 BP.
XX
AC AAZ25728;
XX
DT 05-JAN-2000 (first entry)
XX
DE Stachybotrys chartarum phenol oxidising enzyme genomic DNA.
XX
KW Stachybotrys chartarum; phenol oxidising enzyme; colour; dye;
KM detergent; anti-dye transfer; stain removal; bleaching; ds.
XX
OS Stachybotrys chartarum.
XX
PN WO949010-A2.
XX
PD 30-SEP-1999.
XX
PF 23-MAR-1999: 99WC-EP02042.
XX
PR 24-MAR-1998: 98US-0046969.
PR 22-DEC-1998: 98US-0218702.
XX
PA (UNIL ) UNILEVER NV.
PA (UNIL ) UNILEVER PLC.
PI Convents D, Amory A, Wang H, Dhaese P, Lambrechts-Rongvaux A;
PI Wang C;
DR WPI: 1999-601211/51.
DR P-PSDB; AAY45222.
XX
PT Detergent composition containing phenol oxidase from Stachybotrys, used
XX to bleach stains and prevent dye transfer
XX
PS Example 14: Fig 6; 56pp; English.
XX
CC The present invention describes a detergent composition containing a

```


PN WO949020-A2.
 XX 30-SEP-1999.
 XX 23-MAR-1999; 99WO-US06327.
 PF 24-MAR-1998; 98US-0046969.
 PR 22-DEC-1998; 98US-0218702.
 PR 22-MAR-1999; 99US-0273957.
 XX (GENEM) GENENCOR INT INC.
 XX Amory A, Wang H, Dhase P, Lambrechts-Rongvaux A, Wang C;
 PI WPI: 1999-591088/50.
 DR P-PSDB: AAY39992.
 XX Novel enzyme for modifying coloured compounds used to prevent
 PT dye-transfer.
 XX Claim 21: Fig 6: 64pp: English.
 XX This sequence encodes the Stachybotrys chartarum phenol oxidase enzyme
 CC of the invention. The invention is used to modify a coloured compound and
 CC prevent dye transfer during fabric washing, or for stain bleaching or
 CC anti-dye transfer. It is useful in the detergent, paper and pulp, textile
 CC and food industries.
 SQ Sequence 3677 BP: 822 A; 1056 C; 849 G; 947 T; 3 other:

Query Match 15.3%; Score 300.4; DB 20; Length 3677;
 Best Local Similarity 55.2%; Pred. No. 3.2e-72;
 Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;

QY 316 ctctctccacagatctaccctgatactgagccgcgcaaatggttgataagatggcat 375
 DB 1491 ctgaccccttcagatctaccacaccccttcgcctgcacactctgcgcgaagatggcat 1550
 QY 376 gtccacagagacatcatcatctctctctgagcagtgatgttcgcctcttgaa 435
 DB 1551 gtagcctgtctcactcttcaatgctccagaggaacagagactgtatgtgttcaaa 1610
 QY 436 cagcggagagaaacactctcccaacagcgttcaacttgcagcgtcttctctgagctcc 495
 DB 1611 ca-----atgcccacgctgagaaactcgtccatctgcacagcgtcccccagtgctccc 1664
 QY 496 ctttgatgttggtgagagacactcccaagcctgcgagtagaagattactactacc 555
 DB 1665 ctctgactgttggtgagagatgtgaccttccctgcgagtagaagattactacttcc 1724
 QY 556 caacagagagcgtccgcgcatgtctgtacatgacacatgc-----597
 DB 1725 caactacacatccgcgcgctctctgtgataccttgacaacgcttcatgaagtatgcta 1784
 QY 598 -----catgtccatcacgcgcgag 616
 DB 1785 cgagccttactcttcttgctactccttgctaaacacttctctgtagactgtag 1844
 QY 617 aacgcctataatggtcagcgtgtgtctataatgataccagaccggtctgagagtgcccg 676
 DB 1845 aatgctactcttgtagcagcgtgcctacatataacagcagagctgagagtgctctc 1904
 QY 677 aacctccacagcgcgtagcagagttgatatcccttgcttgatctgacacagatcac 736
 DB 1905 ggtctctcctcagtgagcatatgagatgcgataatccctctgactcctgcgcgcaagatcat 1964
 QY 737 aacgcagagcagcatctctctctccacaaatgagagagttccacactctctgagagtgag 796
 DB 1965 aagcgcagatgtatcctctgctgcagcagaggtgagagacaggaactctgtgagagtgctc 2024
 QY 797 attcaagtgtagaagttgagcccatgtgatgtctcagatccctcagaagatcagatgtatga 856

DB 2025 atccatgtc-----2033
 QY 857 aattgtcagatctctaacagatgctatcaacagacgtcagccttgagctatgctaacg 916
 DB 2034 -----aagcgaacgacatgagcttcttcttaag 2061
 QY 917 tgcagcgcgcaagtaacgccttcgcctctcaacgctgcgtctcaagctcttctgcctc 976
 DB 2062 tccagcccgcaagtaacgcttctcgaatctcccaacgctgcgtctctgtctgtctcctc 2121
 QY 977 tgaatcttgctactcttgtagagatctcagagcagcagctctccctcagcagctatgcgcgtg 1036
 DB 2122 tctaacctcgtcagagaccagctctcccaacgtccagaatcccttccaaagatctgctctg 2181
 QY 1037 acggtgtctcgtcttgagggccctgttgaaactctaacactctgatacatctatgagccagc 1096
 DB 2182 atgtcgtctcctcttaagcccccgttcagaccccttaacctctatctgtctgtccgagc 2241
 QY 1097 gctggaagtgta-----1111
 DB 2242 gtaacgagatcatattgtatgctccctccctcacaagatgagtcagaactctaagac 2301
 QY 1112 -----tcgactctccactctgcgtgcgcagctcatgctatccgcaactctcgt 1162
 DB 2302 taacacttgtagactcttaacaaacttgcgtgcagactcttgactctgcgcgaacgttgcgt 2361
 QY 1163 gtgctacggtctcgtgtgtgagcctgagttgatacaactctgacaaggtcatgctgatactg 1222
 DB 2362 agacccaagatgtcgcgcagagagatgagtagctctgcacactctcagggatgctgcttg 2421
 QY 1223 tegtatgaaagctcttgagctgcgcgaacactctgaagtgctcctcgaacactcgaagatg 1282
 DB 2422 tctgacgtctgcagactgtgag---gacaacagcagatgcctccactctccgtcagagc 2478
 QY 1283 tttctttcccgagggcggaactgtagaccgcgaaccccaactgtatgtagaacttca 1342
 DB 2479 tttctttccct-----ctcaagaagaagcccgccgcgaacgaactca 2523
 QY 1343 cctctgcgcgtgttaatgtagacagtgtagacaactcaacaggaattactctctcgtgaatgtaga 1402
 DB 2524 agtttgaaagcagcaacgagacactactctgaacagatgtgtgtctgcgcagatgaag 2583
 QY 1403 accgtgtctcgcgaatgtgccccgcgaacactgtgtagatctgagcacttgagaacact 1462
 DB 2584 agcgtgtctcgcgaagccgagcagctgcgacagctltgaggtctgtagcgcgaagaaactct 2643
 QY 1463 ccaagcttgtagaactccctgttcaatcaactctgttactcgcagagctcttccgtt 1522
 DB 2644 ctgagagctggaagccaccccgctccacatcaactgtgtacttcaagatccctcaagcgaa 2703
 QY 1523 ccactgcgcgtg---agtcgaagccttatgagcgtgctgtctcaagagatgtgtctgagc 1579
 DB 2704 ctggtgtctgtgcagagatcatgctcctaagagatctgtgtcttaagagatgtctgtctg 2763
 QY 1580 tggctcgtctgtaggtgtgtctatgtgtgagcccaactgaactcctctccgtaagttctcg 1639
 DB 2764 tggcagaggggtgagaccctgagccatcagagcccaactcaaaccccttgagc-----2812
 QY 1640 cctttaaactaaatgttttaactcatgtctaaacatctcaagaagtggtgtactaagtgtaga 1699
 DB 2813 -----tggaagcttaactatgtgga 2830
 QY 1700 ctgcacaacactgatacagagacagacagatgagctgtgcttcaatgaactgcttct 1759
 DB 2831 ctgtcaaacctcatctcaagagagatacgaatgatagtgtgtatcttaacatgtaacgcgc 2890
 QY 1760 cgtgtactatgtgatacaactacacgagatcttgaccccatgtgagcctctctgagagcgc 1819
 DB 2891 ggaagagagaggaatatcttcaagaggaactctgagagcccatgaacccaagtgagcgcgc 2950
 QY 1820 ccgcccctctcctctcgtgagagatctgagaaatgagctcgtgagacttcaagcagactgtccat 1879
 DB 2951 cgttcttaacacgcgaacgacttccatgtctgcgtgtgaaacttctccgcgagttcat 3010

QY 1880 cactgacgcacatcagagatgcttagctcaaccctacgcccagctgatga 1933
||||| ||| | ||||| ||||| | || ||| | |||||
Db 3011 cactgcccagctgacagagactgcccagcagcagcctgacacccgctcatga 3064

RESULT 15

AAA50018
ID AAA50018 standard; DNA: 3677 BP.

AC AAA50018;

DT 10-OCT-2000 (first entry)

DE Stachybotrys chartarum phenol oxidising enzyme gene.

XX phenol oxidising enzyme; detergent; bleaching; ds.

OS Stachybotrys chartarum.

XX Key Location/Qualifiers

FT CDS 1044..3095

FT /*tag= a

FT /note= "contains introns"

FT exon 1044..1237

FT /*tag= b

FT /number= 1

FT Intron 1238..1285

FT /*tag= c

FT exon 1286..1324

FT /*tag= d

FT /number= 2

FT Intron 1325..1376

FT /*tag= e

FT exon 1377..1449

FT /*tag= f

FT /number= 3

FT Intron 1450..1503

FT /*tag= g

FT exon 1504..1776

FT /*tag= h

FT /number= 4

FT Intron 1777..1835

FT /*tag= i

FT exon 1836..3095

FT /*tag= j

XX WO200039306-A2.

XX 06-JUL-2000.

XX 20-DEC-1999; 99WO-EP10287.

XX 23-DEC-1998; 98US-0220871.

XX 23-JUN-1999; 99US-0338723.

XX (UNIL.) UNILEVER NV.

XX (UNIL.) UNILEVER PLC.

XX (HIND-) HINDUSTAN LEVER LTD.

XX Bodie EA, Van Der Velden S, De Vries CH, Wang H;

XX WPI: 2000-514528/46.

XX P-PSDB; AAY95537.

XX Detergent composition comprising novel phenol oxidising enzyme obtained

XX from fungus or bacteria, useful for pulp and paper bleaching, bleaching

XX color of stains on fabric and for anti-dye redeposition

XX Claim 1; Fig 1A-B; 45pp; English.

XX The present sequence is that of the Stachybotrys chartarum MUC1 38898

XX phenol oxidising enzyme gene, including promoter and terminator

CC sequences. The gene was isolated from genomic DNA using primers
CC (see AAA50023-24) based on isolated peptides of the enzyme. The
CC gene codes for a 594-amino acid protein (see AAY95537). The invention
CC relates to detergent compositions comprising novel phenol oxidising
CC enzymes that are encoded by nucleic acids capable of hybridising to
CC the present DNA sequence, provided the enzymes are capable of
CC modifying the colour associated with dyes or coloured compounds, and
CC are produced from a bacterium, yeast or fungus (see AAY95538-40). The
CC phenol oxidising enzymes can be used for pulp and paper bleaching,
CC for bleaching the colour of stains on fabric and for anti-dye
CC transfer in detergent and textile applications. They may also be
CC capable of modifying the colour in the absence or presence of an
CC enhancer. Expression vectors and host cells comprising a nucleic
CC acid encoding a phenol oxidising enzyme, methods for producing the
CC phenol oxidising enzyme, and methods for constructing expression
CC hosts are provided.

XX Sequence 3677 BP; 822 A; 1057 C; 849 G; 948 T; 1 other;

SQ Query Match 15.3%; Score 300.4; DB 21; Length 3677;

Best Local Similarity 55.2%; Pred. No. 3.2e-72;

Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;

QY 316 ctctccaccagatctaccctctgacgtgagccgccaatggttgatagatgagcat 375

Db 1491 ctgacccttcacagattaccacccttcgcccacatctcgtgctgacagtgatcat 1550

QY 376 gtcccaagacctacatcatcgtcttcctgtgacatgagatggtgtccgcttgtaa 435

Db 1551 gacccctgtctacttcaatgattcccaagagacagagactgaattgattacaa 1610

QY 436 cagcgagagagaacacccctcccaacagatccacttcacagggctcttcctgagtc 495

Db 1611 ca-----acgcaacgctgagagactcgtctcactgacagcccccacatcgctgcc 1664

QY 496 ctltgattgtggtctgagagacactaccagcctgaggtatagagattactacc 555

Db 1665 ttctgattgtggtctgagagactcctccctgagagagagagattactactcttc 1724

QY 556 caacagagaggtctgcccagatgcttctgacatgacatgc----- 597

Db 1725 caactacaatccgcgcgcctctgtgtacatgacacagccttcataagagtatgcta 1784

QY 598 -----catgtccatccgcgcgag 616

Db 1785 cgaagccttattctctgtgactccttggtaaccaactccttctgtagacgtcgag 1844

QY 617 aacgctacatggttcagagctggtgtctacatgataccaagaccgctgagatgcctg 676

Db 1845 aatgctacttgttcaagctgctgagcctacatataccaagcagagctgagctctc 1904

QY 677 aacctcccaagcgtcagcagagatttgatatacccttggtctcgaactgcaagcatalac 736

Db 1905 gttcttcctagtgctatgagagattcgatatacccttgatcctgacagcgaagtaact 1964

QY 737 aacgcagagcgaactctctctccaccaatgagaggtttccagctctctgaggtgagact 796

Db 1965 aacgcagatggttacccgctgcttcgacagaggtgagagcagagcctgtggagagatgc 2024

QY 797 attcaatggttaagtttgaccatgtgagatgcttcagatccatagatlatogatgata 856

Db 2025 atccatgtc----- 2033

QY 857 aattgtcatgtcttaaacagtgctctacacagaaggtcagccttgctatgctcaag 916

Db 2034 -----aacgacagacatgaccttccttaag 2061

QY 917 tgcagccggaagatcagccttccttcacagcctgcttcagccttcctgc 976

Db 2062 tccagcccgcaagatcagccttccttcacagcctgcttcagccttcctgc 2121

QY 977 tgaattctgactacttgagatccagagaccagactcccttcaggtcattgcgctg 1036

Db 2122 tctaccctgcagacagcctccccaacg tcaagaattcccttccaag tcatctgcccctg 2181
QY 1037 acggtgtctgcgttgaggcgccgttgtaacactgaacactctgtacatctcatgtgcccagc 1096
Db 2182 atgcctgtctcccttcaagcccccctcaagaaccttaaccttaacctgtctgtccgagc 2241
QY 1097 gctggagggttgta----- 1111
Db 2242 gttacgagatcatattgtatgcccctccctctcaagaa tga tgaagacttaagac 2301
QY 1112 -----tgaactctccaactctgctgcccagtcacatagatatccgaactctcgt 1162
Db 2302 taacactgttagaacttcaaccaacttctgtgcccagactcttgaactctgcgaacgttgcgtg 2361
QY 1163 gtgtgacaggtctgcgtgtgttgagccctgaagtttataaacttgacaagagtcatactgac 1222
Db 2362 agaaccacagatctgcgagacgaagatgagtaacgtctgcactctcgaagtgatgcgtctcg 2421
QY 1223 tctgtatgaagtccttgagtcgcgccgacacttctgaagtgctctgcgaacctcgaagatg 1282
Db 2422 tctgtacagctctgcactgttgag---gacaacagccaggtccctctcaactctcgtgagc 2478
QY 1283 ttccttcccccgaaggcggaacttggagcccccgaaccccaactgtatgacgagacttca 1342
Db 2479 ttccttccct-----cctcacaggaaggcccccgcgcgaacagacttca 2523
QY 1343 ccttcgcccgtgtcta tga tgaagtgacaacacgaagttacactctcgtgagatgcgaga 1402
Db 2524 agtttgaacgcacacacgaacgaactaactctgaacagatgttgcgtcttgcgaatgcaatg 2583
QY 1403 accgctgtctgcgaatgtgcccgcgacacgttgaagatctgcgacttgagaacaact 1462
Db 2584 agcgtgtcttcgagcccgagctcgacacgcttgaggtctggagctcgagactcct 2643
QY 1463 ccaacggttggaactcaacctgttcaacttgaactctgtgactccgaagtccttccgt 1522
Db 2644 ctggaggctggagccaccccgctcaacttgaacttgaactcaagatctcaagcga 2703
QY 1523 ccaactgcgcgtg---agtcgaacctatgagctgtgctcgaagatgttgcgtgc 1579
Db 2704 ctggtgtgtgtgcgacagatcatctccttaacgagctgtgtcttaaggaatgtcgtgcgt 2763
QY 1580 tggctcgtctgaggtgtgtatgttgagggccactaagctccttccgttaagttctcg 1639
Db 2764 tgggcaggggtgagaccctgaacatcagagccactaacaaccttgac----- 2812
QY 1640 cctttacactgaactgtttcaactatgctaatactacaagtggtctacatgttga 1699
Db 2813 -----tggagcttaca tgcga 2830
QY 1700 ctgcacaacactgatccagaggaacacgaatgatgtgtcttcaatgtactgttct 1759
Db 2831 ctgtcaaacctcatctcacagagataacgaatgatgtgtcttcaagtcacgcgcac 2890
QY 1760 cgtgtgactatgtctacaactacacgagatcattgacccca tggagcctctctgagggcc 1819
Db 2891 ggagaggaaggaatatcttcagagagacttcgagaccccatgaaccccaagtgcgcgc 2950
QY 1820 ccgccccttccctctcgtggagagtttgagaatgtgtggtgacttcaagacttgcaat 1879
Db 2951 cgttcccttaacacgaacgaacttccatctcgcgtggaaacttctccgcgcgaatccat 3010
QY 1880 cactgaccccatctgagagatgctagcttcaacccctacgcccaagctgatga 1936
Db 3011 cactgcccagtgagagctgagcgcgagcagagagccgttaaacgcgtctgatga 3064

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 17:55:50 ; Search time 3211.21 Seconds
(without alignments)
8229.620 Million cell updates/sec

Title:

US-09-656-640A-1

Perfect score:

1958
1 ggtaccatcaatgatcag.....ccgctgagagtagaccggt 1958

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:

13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters:

27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estbda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: qb_est1:*
10: qb_est2:*
11: qb_htc:*
12: qb_gss:*
13: em_gss_hum:*
14: em_gss_juv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86.4	4.4	600	BE188240	M7ATIG In
2	86.4	4.4	700	BE187716	BITOX In
3	68.6	3.5	664	BE188099	CFC37-R
4	65.6	3.4	848	BE216983	BE216983 HY-CEA001
5	63.4	3.2	880	BF264805	BF264805 HY-CEA001
6	54.6	2.6	465	BC053739	RH22.9-B
7	51.4	2.6	311	AV426401	AV426401
8	50	2.5	580	BE159802	BE159802
9	48.2	2.4	592	BM143324	BM143324
10	47.8	2.4	766	BE26127	56-88 Ste
11	46.2	2.4	286	AZ714471	RPCI-24-1
12	46.2	2.4	599	BF258534	HVSMF001
13	46.2	2.4	669	BB655683	BB655683
14	45.6	2.3	485	AI876848	u35b06.y
15	45.6	2.3	562	BB651243	BB651243
16	45.6	2.3	596	AI574420	AI574420
17	45.6	2.3	600	BB634973	BB634973

18	45.6	2.3	634	AI574382	AI574382
19	45.6	2.3	738	BE918653	BE918653
20	45.6	2.3	762	AI316242	AI316242
21	44.8	2.3	657	BI527303	BI527303
22	44.2	2.3	705	BE50982	BE50982
23	43.8	2.2	433	AW018870	AW018870
24	43	2.2	552	AZ302933	AZ302933
25	42.8	2.2	473	AV633110	AV633110
26	42.8	2.2	542	AV626368	AV626368
27	42.8	2.2	611	AV391222	AV391222
28	42.6	2.2	725	AI317680	AI317680
29	42.4	2.2	533	BE234675	BE234675
30	42.4	2.2	551	BE321929	BE321929
31	42.4	2.2	591	BF789685	BF789685
32	42.4	2.2	697	BE914800	BE914800
33	42.4	2.2	795	BI559124	BI559124
34	42.4	2.2	885	BI558920	BI558920
35	41.8	2.1	667	BE212100	BE212100
36	41.6	2.1	444	BM100862	BM100862
37	41.6	2.1	518	AA182405	AA182405
38	41.2	2.1	339	F01214	F01214
39	41.2	2.1	541	BE255913	BE255913
40	41.2	2.1	722	BE250985	BE250985
41	41.2	2.1	828	BE749472	BE749472
42	41.2	2.1	1128	AF094955	AF094955
43	40.8	2.1	518	BE031325	BE031325
44	40.8	2.1	534	BM421598	BM421598
45	40.6	2.1	316	BE725648	BE725648

ALIGNMENTS

RESULT 1
BE188240
LOCUS BE188240
DEFINITION M7ATIG In vitro expressed cDNAs Cladosporium fulvum linear EST 25-SEP-2000
Bilirubin oxidase precursor, mRNA sequence.
ACCESSION BE188240
VERSION BE188240.1
KEYWORDS GI:8667479
SOURCE Cladosporium fulvum
ORGANISM Cladosporium fulvum
EST.
Cladosporium fulvum
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothryomycetes Incertae sedis; Mycosphaerellaceae; mitosporic
Mycosphaerellaceae; Cladosporium.
REFERENCE 1 (bases 1 to 600)
AUTHORS Clark,A.J., Rasmussen,S.W. and Oliver,R.P.
TITLE In vitro expressed genes of Cladosporium fulvum
JOURNAL Unpublished (2000)
COMMENT Contact: R.P.Oliver
Neurotrophic Phytopathology Research Centre
Murdoch University
SAB, Perth 6150, Western Australia
Tel: +61-8-9360-7404
Fax: +61-8-9360-6303
Email: roliver@central.murdoch.edu.au
High quality sequence stop: 600
POLYA=NO.

FEATURES

source

1. 600
/organism="Cladosporium fulvum"
/strain="Race 4"
/db_xref="taxon:5499"
/clone_id="In vitro expressed cDNAs"
/tissue_type="Mycelium"
/note="Vector: Lambda bluescript; A mixture of mycelial cultures grown in liquid B5 for 48 hours and transferred for 24 hours to media lacking carbon, nitrogen or supplemented with hydrogen peroxide."
BASE COUNT 131 a 176 c 162 g 128 t 3 others
ORIGIN

/db_xref="taxon:5499"
/clone_lib="In vitro expressed cDNAs"
/tissue_type="Myxcelium"
/note="Vector: Lambda bluescript: A mixture of mycelial
cultures grown in liquid B5 for 48 hours and transferred
for 24 hours to media lacking carbon, nitrogen or
supplemented with hydrogen peroxide"

BASE COUNT 149 a 174 c 174 g 140 t 27 others
ORIGIN

Query Match 3.58; Score 68.6; DB 9; Length 664;
Best Local Similarity 58.1%; Pred. No. 4.4e-06;
Matches 158; Conservative 0; Mismatches 107; Indels 7; Gaps 2;
QY 888 gaacggtcagccttgatgctatgctcaagtcgagcgccgagaaatcagcttccttct 947
DB 57 GAACGTGCAAGCAGGCGCATATCTGAACGTCGAGCCTCGCAAGTACCGATTCAAGCTTT 116
QY 948 caacgtcgcgtctcagcgtcttcgtctgtatcttgctactctgaagattcagagac 1007
DB 117 TGATATGCTCTTAAGAGAGATCTTACGACATCTACATTG-----CTGATGAGAACACAA 170
QY 1008 cgaactcctccctcagcagtcagtcgagtcgagtcgtctgtgaagggccctgtgaac 1067
DB 171 CCCAATCCAAATTCAGAGTATGCTCAGACTCCGGCCTATTGCGCGCTNCGCTGACGCG 230
QY 1068 tgacactctgactctcatalgacgcgctgagagtggtatgactctcccaactt 1127
DB 231 CAAGACGTGGTATTTCATTCAGGAGAGCGCT-TGAGATCATTTAGACTTTANCTGTTT 289
QY 1128 cgtcgcagcagtcacatcagatccgcaacttc 1159
DB 290 CCGCGGCCAGAAATTTACTATGTTTCANCGGTC 321

RESULT 4
LOCUS BE216983 848 bp mRNA linear EST 03-JUL-2000
DEFINITION EST0376 Triticum aestivum Lambda Zap Triticum aestivum cDNA clone
JAL5A_A06_T3 5', mRNA sequence.
ACCESSION BE216983
VERSION BE216983.1 GI:8904523
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 848)
AUTHORS Anderson, J.M., Williams, C.E. and Goodwin, S.B.
TITLE Analysis of an EST database reveals a probable Cf2 resistance gene
homolog in wheat
JOURNAL Unpublished (2000)
COMMENT Crop Production & Pest Control Research Unit
USDA-ARS
1150 Lilly Hall, West Lafayette, IN 47907, USA
Tel: 765-494-5565
Fax: 765-496-2926
Email: janderson@purdue.edu
Seq primer: T3
High quality sequence stop: 848.
FEATURES
source
Location/Qualifiers
1..848
/organism="Triticum aestivum"
/strain="P29"
/db_xref="taxon:4565"
/clone="JAL5A_A06_T3"
/clone_lib="Triticum aestivum Lambda Zap"
/tissue_type="leaf"
/dev_stage="9 day old seedlings"
139 a 312 c 263 g 134 t

ORIGIN

Query Match 3.48; Score 65.6; DB 9; Length 848;
Best Local Similarity 55.2%; Pred. No. 2.7e-05;
Matches 128; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 526 gctcgcagatcaagaattactactaccacacagcagcgtcgcgcagcttgtagta 585
DB 1 GCACGAGGACTCGCCGCGGCTACGAGTACCGAACACACAGCTCTCCGTAACCTGGA 60
QY 586 ccatgacatgcatgtccatccacgcgcgagaaagcctacatggtgaagctgtgtcta 645
DB 61 CCAGACACACGCGATGGCGCTCACCGCGCTCAACATCTCGTGGCTCTCGCGCGGTA 120
QY 646 catgattccagacccggtcgtgagatgctcctgaacctccacagcgctgaaggttga 705
DB 121 CCGTGTCCGCGACCCGACCCGACCGGCCCTCGGCTCCCTCCGCGCGGAGTTGCA 180
QY 706 tatcccttggtctgactcgaagcgatcacgcagcagcactctctc 757
DB 181 CCGGACCTGCTCTCTTTCAGCCGCACTTCAAGAGCGAGCGCGCTCTTC 232

RESULT 5
LOCUS BF264805 880 bp mRNA linear EST 23-OCT-2001
DEFINITION HV_CEA0010619f Hordeum vulgare seedling green leaf EST library
HVCDNA0004 (Blumeria challenged) Hordeum vulgare cDNA clone
HV_CEA0010619f, mRNA sequence.
ACCESSION BF264805
VERSION BF264805.2 GI:13261741
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 880)
AUTHORS Wing, R., Close, T.J., Kleinbols, A., Wise, R., Wei, F., Begum, D.,
Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi,
D.W., Fenton, R.D., Oates, R. and Main, D.
TITLE Development of a genetically and physically anchored EST resource
for barley genomics: Blumeria infected incompatible (Mia13)
seedling leaf cDNA library
JOURNAL Unpublished (2001)
COMMENT On Nov 17, 2000 this sequence version replaced gi:11195799.
CONTACT: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Total hg bases = 260
Seq primer: AATTAACTCTCACTAAAGCG
High quality sequence stop: 798.
FEATURES
source
Location/Qualifiers
1..880
/organism="Hordeum vulgare"
/cultivar="C116155 (Mia13)"
/db_xref="taxon:4513"
/clone="HV_CEA0010619f"
/clone_lib="Hordeum vulgare seedling green leaf EST
library HVCDNA0004 (Blumeria challenged)"
/tissue_type="seedling green leaf"
/lab_host="TJ121"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
C.I. 16155 (Mia13) plants were greenhouse grown in the R
wise lab at Iowa State University, Ames, IA; 7 day old
green seedlings were challenged with isolate A27 (AVRMia13
of Blumeria graminis f. sp. hordei, and leaves were
harvested 20 and 24 hr post-inoculation and snap frozen;

uninoculated leaves were harvested 20 hr post-inoculation (Mel, Wise). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one cDNA library was made, and 1 million pfu were in vivo excised to give pluscript SR(-) cDNA phagemids (Choi, Close). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wang). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wang, Yu, Frisch, Henry, Simmons, Gates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wang R, Kleinholz A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. *Barley Genetics Newsletter* 31:29-30. (<http://wheat.pw.usda.gov/gppages/bgn/31/cover.html>)"

	Query Match	2.6%	Score 51.4	DB 9	Length 311:
	Best Local Similarity	52.0%	Pred. No. 0.044:		
	Matches 115,	Conservative 0;	Mismatches 106;	Indels 0;	Gaps 0:
Oy	537	acaaaggattactactaccaccaagcgcaagcgtctccgcgatgcttgtagtaccatgacaag	596		
Db	4	AGAAAGCCTTACACATGCCAACCAACAACCAACCGAATCTATTGATCCATTGACCATG	63		
Oy	597	cgaatgccataccagcccagaagagcgtacatagggtaacgcgtgtctcaatgatgaccaa	656		
Db	64	CCATGGGGTTGACCCGGGTCAACTCTTCGTGCCTTAATAGGACCTTACATCATTCGCC	123		
Oy	657	aaccggtgtagagatgccccitgaacctcccccaagcygtaacygcaggtttatatcccttgg	716		
Db	124	ACCCTTCATCGAGACCCCGCTTGGAATTACCCATATGGAGACGACTGATGACCCGTTGG	183		
Oy	717	tcttcgactgcacaagagatatacaagcagcggaactcttc	757		
Db	184	TCTTGTTCGATCGTAGCTTCCTCCGAGACGAGGCTTCATCATTCAC	224		

RESULT	8
BJ159802/c	
LOCUS	BJ159602 580 bp mRNA linear EST 24-JAN-2002
DEFINITION	GjBt59602 full length cDNA library; chloromemata and young gametophores Physcomitrella patens subsp. patens CDNA clone phi17c15 5', mRNA sequence.

ACCESSION	BJ159802
VERSION	BJ159802.1
KEYWORDS	GI:18327798
SOURCE	EST.
ORGANISM	Physcomitrella patens subsp. patens.
REFERENCE	Physcomitrella patens subsp. patens.
AUTHORS	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Bryophyta: Bryopsida: Funariidae: Funariales: Funariaceae: Physcomitrella. 1 (bases 1 to 580)
TITLE	Fujita, T., Shin-I, T., Seki, M., Kamiya, A., Uchiyama, I., Nishiyama, T., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y. and Hasebe, M.
JOURNAL	Comparison of the moss Physcomitrella patens genome with flowering plants genome
COMMENT	Unpublished (2002)
	Contact: Tadasu Shin-I

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel.: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshinigenes.nig.ac.jp

A backbone of the vector is basically from pBluescript(KS), that was in vivo excised from a modified lPS phage vector (Mo bi Tec, Germany). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector. cDNA insert could be amplified with conventional 7F and 73 primers. This full-length cDNA library was generated basically according to the method described in The Plant J 15, 707-720 (1998) Seki M. et al. Protoplasts were blended by the POLYTRON, and then cultivated on the BODiG medium for 13-14 days under the continuous light.

FEATURES	Location/Qualifiers
source	1. .580

```

/organism="Physcomitrella patens subsp. patens" *
/db_xref="taxon:145481"
/clone="pph17c15"
/clone_lib="full length cDNA library, chloronemata and
young gametophores"
/tissue_type="mixture of chloronemata and young
gametophores with 2 to 5 leaves"
BASE COUNT      124 a      138 c      170 g      148 t
ORIGIN

```

Query Match 2.68; Score 50; DB 10; Length 580;

	Best Local Similarity	51.0%;	Pred. No.	0.14:	
	Matches	156;	Conservative	0;	Mismatches 135; Indels 15; Gaps 1;
OY	888	gaagcgtcagcccttggcgtctatgcttaacgctgcagccgcgcaaataccgcttcgcctcct	947		
Db	445	GAACCGCAAAACATGCGCTTACCTGCTGCTCAAGGCTCGCAAGTACCGCTTCCGCTCTCT	386		
OY	948	caacgctccgctctcaagctctcttcgcctctgatacttgtaacctgagatcagagac	1007		
Db	385	GAATGTGTGGCAATGGAAGATTTTGGAGCTTTCTGTGAGACCCCA-----	339		
OY	1008	cagacttccctccaagctcaatctgcccgcgtacagcgtgctcgtcttgtagggccctgttgacac	1067		
Db	340	CAAAACCGCGCTTCATTCAAATTTGCTACCGACGAGACGACTCCGAAAAGCCGATGTATGT	281		
OY	1068	tgacacctctgacatcctatctatgcccagcgcgcggaggttgtatcagctctccacct	1127		
Db	280	GAGCAGCTTAAATCCGTGCTCCACGCGGAAGGCTGAGCTTCATCAATGATTTTCCCTTGT	221		
OY	1128	cgcctggccagctcacgcatalccgacaaaccttcgttgctgacggtctctcgtgttgagcc	1187		
Db	220	CGCGCGGGAAGCATTTATTTCCAAAAACTCAGACACTTGTGCGCCCTATCCCGAGGGGGAATCC	161		
OY	1188	tgagct	1193		
Db	160	AGCGTT	155		

RESULT	9
BMI43324	
LOCUS	BMI43324
DEFINITION	BMI43324 592 bp mRNA linear EST_29-NOV-2001
ACCESSION	S6142210.y1 Gm-C1072 Glycine max CDNA clone SOYBEAN CLONE ID:
VERSION	Gm-C1072.2563 5' similar to Tr:O23123 O23123 F19G10.5 PROTEIN.
KEYWORDS	mRNA sequence.
ORGANISM	BMI43324
SOURCE	BMI43324.1 GI:17153391
GLYCINE	EST.
SOYBEAN	soybean.
PROTEIN	glycine max

REFERENCE	1 (pages 1 to 592)
AUTHORS	Shoemaker, R., Keim, P., Vockin, L., Eppelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Knudsen, T., Martin, J., Beck, C., Wyle, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.
TITLE	Public Soybean EST Project
JOURNAL	Unpublished (1999)
COMMENT	Contact: Shoemaker R/Public Soybean EST project

Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp., 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: cust@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Glbco
High quality sequence stop: 423.
Location/Qualifiers
1. .592

```
FEATURES
source
location/Qualifiers
1..592
/organism="Glycine max"
/db_xref="taxon:3847"
/clone-"SOYBEAN CLONE ID: Gm-cl072-2563"
/clone_id="Gm-cl072"
/tissue_type="seedlings induced for symptoms of SDS
(Sudden Death Syndrome) disease"
```


COMMENT

Other GSSS: RPCI-24/138E2.TV
Contract: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@email.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi1.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tcdb/bac_ends/mouse/bac_end_intro.html
Plate: 138 row: E column: 2
Seq primer: SP6
Class: BAC ends.

FEATURES	Location/Qualifiers
source	1. .286

```

/organism="Mus musculus"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-138E2"
/clone_lib="RPCI-24"
/sex="Male"

```

/cell_type="Spleen/Brain"
/note="Vector: pPARBAC1, Site_1: BamH1, Site_2: BamH1;
RPCT-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pPARBAC1 cloning vector at the
BamH1 sites using MboI partially digested male C57BL/6J
DNA."

BASE COUNT ORIGIN	71 a	79 c	67 g	69 t
----------------------	------	------	------	------

ORIGIN

Query Match	2.48;	Score 46.2;	DB 12;	Length 286;
Best Local Similarity	52.38;	Pred. NO. 0.78;		

479 tcttctctcgagctcccttctgatgtgtgctgagacatacccaagctggcgagtac 538

Db 36 TCTTCTCTTAGTGGCAGGAGATGCCAAGATAAGACATCTCCTGGTCTCTCCCAAT 95

539 aaggattactactacccaacagcagctgccgcgcatgtcttgytaacatgacatgcc 598

Db 96 GGGGAGAGCTGAGCCCAACACGACGCGGATCTCAGTGGTGTGGAATGATGACTTCTTC 155

QY 599 atgtccatcaccgcgcgagaacgcctacatggtggtcagcgtgtgtctacatgatccaagac 658

Db 156 TTTACCCCTCACCATCTACACAGCCACACATCGAGGATGCCGGCATATACAAAGTGTTGGTC 215

QY 659 ccgctgagatgcc 673

216 ACGGCTGAGGACGGC 230

RESULT 12
BF258534

LOCUS	599 bp	mRNA	linear	EST	22-OCT-2001
BF258534					
DEFINITION	HV5HEF0015P21f Hordeum vulgare seedling root EST library HVCNNA00007				

HVSMEf0015P21f, mRNA sequence.

ACCESSION	BF258534
VERSION	BF258534.2
	GI:13119477

KEYWORDS
Cognitive

SOURCE ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Dicotyledonae; Rosales; Rosaceae; Prunella

Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, Pooidae
; Triticeae; Hordeum.

REFERENCE	1 (bases 1 to 599)
Wing, R. A., Close, T. J., Kleinbols, A., Wise, P., Beum, D., Erisch, D., and	

Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton

R.D., Oates, R. and Main, D.

TITLE

**JOURNAL
COMMENT**

COMMENTS

Development of a genetically and physically anchored EST resource
for barley genomics: Morex unstressed seedling root cDNA library
Unpublished (2001)
On Nov 16, 2000 this sequence version replaced gi:11187647.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total bp bases = 451
Seq primer: AATTACCTCCTCAAGGC
High quality sequence stop: 552.

FEATURES

```

/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HWSMf001521f"
/clone_1fp="Hordeum vulgare seedling root EST library
HVCMDNA007 (Etilated and unstressed)"
HVCMDNA007 "Etilated root"
HVCMDNA007 "Etilated root"

```

```

/rabuse-typer-sealing tool
/lab_host="TYC121"
/noise=Vector: lambdaZAP: site_1: EcoRI: site_2: XhoI:
conditions In the dark at room temperature on filter paper
with water, nystatin and cefotaxime in covered
crystallization dishes. Five-day old seedling roots were

```

BASE COUNT	96 a	215 c	206 g	82 t
IRIGIN				

Query Match	2.48; Score 46.2; DB 10; Length 599
-------------	-------------------------------------

Matches 105; Conservative 0; Mismatches 78; Indels 2; Gaps

889 aacgctcagccttgcgtactatgcctcaacgtgcagccgcgcaagtacgcgttcgcgtcttc 948

OCT 01 1967

U.S. DEPARTMENT OF AGRICULTURE
LIBRARY

NOV 1 1967

U.S. DEPARTMENT OF AGRICULTURE
LIBRARY

949 aacgctgcgctctaacgctctttcgtctgtatctgtacct--ctgagattcagaga 1006

b 137 AACGCCAGCAACGCGCTTCTTCCGCTCTCGCTCTCCGGCGGCTTGTGCAC 196

1007 ccaagcttccttccaagatcatatccagctaacatgattctacttaaaagaccctattgaca 1066

[illegible]

19/ GTCGCTCCGACATCCGTGTAACCTCGCCCGGCGGTGCGCCACGAGCGGTTCCTGCTCGCG 256


```

/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
(CACATGCTG); Site_2: DraIII (CACCATGCTG); 1st strand cDNA
was primed with an oligo(dT) primer
TAAGGCGCTTTTCTTTTTTTTTTTTTT; double-stranded cDNA was
ligated to a DraIII adaptor (TGGTGGCCATGCTG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACATGCTG, 3' site CACCATGCTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo,
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGGCTCTTAAAGCTGG and 3' end
primer CGACCTCAGCTCGACGACA."
122 a 137 c 120 g 106 t

```

Query Match	2.3%	Score 45.6;	DB 9;	Length 485;
Best Local Similarity	52.0%;	Pred. NO. 1.5;		
Matches 102;	Conservative 0;	Mismatches 94;	Indels 0;	Gaps 0

Oy	478	ctctctctcgcagacccctctttagtcttgccgaagaacctaccagcctggcagta	537
Db	265	CTTCTCTGTGTAAGTGGCAGGAGATGCCAAAGATTAAGGACATCTCTCTGTTCTCCCCAA	322
Oy	538	caagatattactactaccaccaacagcagcagctgcgccgaatgctcttgtaaccatagcacatgc	597
Db	325	TGGGAGAGAGCTGAAGCCCAACACAGACGGGATCTCAGTGGTGTGGAAATGATGACGACTC	384
Oy	568	catgtcacaatccgcgcggaagaagcctcataccttgatcagccttgctgctacatacga	657
Db	385	CTCTACCTCTCAACCATCTTCAACAGCGCAACATGACAGATGCGCGCATATACAAGTGTGGT	444
Oy	658	ccgcgcctgagatgctc	673
Db	445	CACGGCTGAGGAGCGC	460

RESULT	15
BB651243	
LOCUS	BB651243
DEFINITION	BB651243 562 bp mRNA linear EST 26-OCT-2001 BB651243 RIKEN full-length enriched, 0 day neonate cerebellum Mus musculus cDNA C230070H21 5', mRNA sequence.

VERSION	BB651243.1	GI:16485497
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE
AUTHORS
Akakawa, T., Carnicci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, H.,
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
1 (bases 1 to 562)

TITLE
JOURNAL
COMMENT

Unpublished (2001)
Contact: Yoshihide Hayashizaki

Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., K.
, M., Koya, S., Matsuyama, T., Miyazaki, K., Nomura, K., Ohno, M.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sa.
, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takada, Y., Tanaka, T., Toya,
Morimatsu, M., and Hayashizaki, Y.

RIKEN MOUSE ESTS (Aikawa, T., et al. 2001)

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9222
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Yamamoto, T., Shibata, Y., Hayashi, N., Sugahara, Y., Shibata, K., Itoh,
Yamamoto, T., Shibata, Y., Hayashi, N., Sugahara, Y., Shibata, K., Itoh,

M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapped-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-285 (2001)

Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamana, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome*, 12, 673-677 (2001)

Please visit our web site (<http://genome.riken.go.jp>) for further details.

e mouse tissues.

FEATURES
source

```
BASE COUNT      141 a      160 c      142 g      119 t
ORIGIN
/organism="Mus musculus"
/db_xref="taxon:10090"
/cclone="C230070H21"
/cclone_1lb="RIKEN full-length enriched, 0 day neonate cerebellum"
/tissue_type="cerebellum"
/dev_stage="0 day neonate"
/lab_host="DH10a"
/note="Site_1: Salt; Site_2: BamHI; CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGGAGAGAGATCCACGAGCTCTTTTTCCTTTTTTTTNN 3'], cDNA was prepared by using reverse transcriptase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 479.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGAGATTCGCAGGTAAATAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PLC I."
```

Query Match	2.3%	Score 45.6;	DB 9;	Length 562;
Best Local Similarity	52.0%;	Pred. No. 1.6;		
Matches 102; Conservative	0;	Mismatches 94;	Indels 0;	Gaps

Qy	478	ctcttctctcgaagctcccttttggttggtgcgaagaacataccagcctgtagagta	533
Db	303	CTTCTGTGTAAAGTGGCAGAGATGCCAAGATAAAGACATCTCTGTTCTCCCAA	362
Qy	558	caagatatactactatacccaacagcagcgtctgcgcgatgtttgltacatgacatagc	597
Db	363	TGGGGAGAGCTGAGCCCAACCCAGCCAGGATCTCATGTGTGTGAATGATGACGATC	422
Qy	598	catgccaatcaacgcgcgaagaacgctcatatggtgcagcgttggttctacatgataccgga	655
Db	423	CTTACCTTCAACCATCTTACAAAGCCAAACATGACAGATGCCGCAATACAAAGTGGTGGC	482
Qy	658	cccgactgagatgccc	673

Mon Oct 7 11:28:38 2002

us-09-656-640a-1.rst

Page 10

Db 483 CACGCTGAGGACGGC 498

Search completed: October 5, 2002, 18:54:14
Job time: 3504 sec

QY	1	ggtacatcaacatgatacgaacgaagatctggaacccgtgtgcctcgggacctgtgcgtatc	60
Db	1	ggatccatcaacaagatacgaacgaagctatccgaacccgtgtgtcccgcccttgcgtgtatc	60
QY	61	ggcgacgactctgtctgaatgcagataccgtttctgtctgatctgacagacatatccttcggt	120
Db	61	ggcgacgactctgtctgaatgcagataccgtttctgtctgatctgacagacatatccttcggt	120
QY	121	ctcaccaagagcagacagcagctgtagctctccctctggccttgtatcgaagtgtcctctgcg	180
Db	121	ctcaccaagagcagacagcagctgtagctctccctctggccttgtatcgaagtgtcctctgcg	180
QY	181	atccctctctctgaaagcgccccaaatagataatctataagctctgcagagcccaagct	240
Db	181	atccctctctctgaaagcgccccaaatagataatctataagctctgcagagcccaagct	240
QY	241	tgtcatcatctgcagatccggtcccccacccccaacactggaagagagacatctgtgactcga	300
Db	241	tgtcatcatctgcagatccggtcccccacccccaacactggaagagagacatctgtgactcga	300
QY	301	gatgagagattagagacctctctcccaacagatctaacctgcatcttggaagccgagccaacatcgt	360
Db	301	gatgagagattagagacctctctcccaacagatctaacctgcatcttggaagccgagccaacatcgt	360
QY	361	tgtgataagatgagcatgtcccccagacactacatcatatcgttccctctgtgacatgagagt	420
Db	361	tgtgataagatgagcatgtcccccagacactacatcatatcgttccctctgtgacatgagagt	420

QY 421 tgcgcgttcgtgaacagcggagagaacacactctcccaacagcgttcacttgacggctc 480
|||||
Db 421 tgcgcgttcgtgaacagcggagagaacacactctcccaacagcgttcacttgacggctc 480
QY 481 tttctctcgcgcgttcctcttgatgtgtggtctgagagaacactccacagcctgagatcaaa 540
|||||
Db 481 tttctctcgcgcgttcctcttgatgtgtggtctgagagaacactccacagcctgagatcaaa 540
QY 541 ggaattactatcccccaacagcggagcgtgcgcgcgaatgctttgtatcatatcacatgcat 600
|||||
Db 541 ggaattactatcccccaacagcggagcgtgcgcgcgaatgctttgtatcatatcacatgcat 600
QY 601 gtcacatcacgcgcgagagaacgcctacatctgagctggtgtctacatctgacatcaagacc 660
|||||
Db 601 gtcacatcacgcgcgagagaacgcctacatctgagctggtgtctacatctgacatcaagacc 660
QY 661 ggtctgagagatgacctgaactcccccagcgtacagcgaatgtgatctcccttggtctc 720
|||||
Db 661 ggtctgagagatgacctgaactcccccagcgtacagcgaatgtgatctcccttggtctc 720
QY 721 gactgcacaacgatacaacagcagacgacactctctctccacacatgagaggtttccag 780
|||||
Db 721 gactgcacaacgatacaacagcagacgacactctctctccacacatgagaggtttccag 780
QY 781 ctctcggggtgacgttattcaagctgtaagctgagcccatctgagatgcttcagatccctag 840
|||||
Db 781 ctctcggggtgacgttattcaagctgtaagctgagcccatctgagatgcttcagatccctag 840
QY 841 aagatcatcagatgataatgatactgcatctgaacagctgatacagaagacggttcagcct 900
|||||
Db 841 aagatcatcagatgataatgatactgcatctgaacagctgatacagaagacggttcagcct 900
QY 901 tggcctatgctcacaagctgacagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 960
|||||
Db 901 tggcctatgctcacaagctgacagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 960
QY 961 tcaagctctctcgtctgtatctgttactctgagagatcagaagacagactccctc 1020
|||||
Db 961 tcaagctctctcgtctgtatctgttactctgagagatcagaagacagactccctc 1020
QY 1021 caaggctatgctcgc 1080
|||||
Db 1021 caaggctatgctcgc 1080
QY 1081 atctctatgagc 1140
|||||
Db 1081 atctctatgagc 1140
QY 1141 atcgataatccgcaactctctcgtgtctgacggtctcggtgttgagcctgagcttgataac 1200
|||||
Db 1141 atcgataatccgcaactctctcgtgtctgacggtctcggtgttgagcctgagcttgataac 1200
QY 1201 actgacaaggtcagctgacgtctgctgtgagagctcttgagctcgcgcgcgcgcgcgcgcgc 1260
|||||
Db 1201 actgacaaggtcagctgacgtctgctgtgagagctcttgagctcgcgcgcgcgcgcgcgcgc 1260
QY 1261 gttcgc 1320
|||||
Db 1261 gttcgc 1320
QY 1321 cccactgatacagagactctacactctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1380
|||||
Db 1321 cccactgatacagagactctacactctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1380
QY 1381 gttactctctcgc 1440
|||||
Db 1381 gttactctctcgc 1440
QY 1441 atctgagcacttgagacaacactcccaacggttgagctacacactctgacatccactcgcgt 1500
|||||
Db 1441 atctgagcacttgagacaacactcccaacggttgagctacacactctgacatccactcgcgt 1500

QY 1501 gactccagatctctctcgtctccactgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1560
|||||
Db 1501 gactccagatctctctcgtctccactgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1560
QY 1561 cttaagagatgtgtctgagc 1620
|||||
Db 1561 cttaagagatgtgtctgagc 1620
QY 1621 ccttcccgaaatctcgc 1680
|||||
Db 1621 ccttcccgaaatctcgc 1680
QY 1681 gttggtctacatgtgtgacgtccacaacactgataccaagagacacagatgagctg 1740
|||||
Db 1681 gttggtctacatgtgtgacgtccacaacactgataccaagagacacagatgagctg 1740
QY 1741 ctctcaatgctacagctctcgtgtgactatggtctacacactacacagatctatgacccca 1800
|||||
Db 1741 ctctcaatgctacagctctcgtgtgactatggtctacacactacacagatctatgacccca 1800
QY 1801 tggagcctctctgagagccgc 1860
|||||
Db 1801 tggagcctctctgagagccgc 1860
QY 1861 acttcagagagcttgccatctacatgacccgaatcagaagatgctgacttaccctcag 1920
|||||
Db 1861 acttcagagagcttgccatctacatgacccgaatcagaagatgctgacttaccctcag 1920
QY 1921 cccagcgtatgatagtgc 1958
|||||
Db 1921 cccagcgtatgatagtgc 1958

RESULT 2
US-09-401-476-3
; Sequence 3. Application US/09401476
; Patent No. 6168936
; GENERAL INFORMATION:
; APPLICANT: Wang, Huaming
; TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes
; FILE REFERENCE: GCS84
; CURRENT APPLICATION NUMBER: US/09/401,476
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Stachybotrys chararum
US-09-401-476-3

Query Match 99.4%; Score 1946.2; DB 4; Length 2095;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1948; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 atccatcaaatgatcagccaagctatcgc 62
|||||
Db 67 aagatcaaatgatcagccaagctatcgc 126
QY 63 cggcagcctctgctgacgacatcgcgtctgctgacatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 122
|||||
Db 127 cggcagcctctgctgacgacatcgcgtctgctgacatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 186
QY 123 caccagaagc 182
|||||
Db 187 caccagaagc 246
QY 183 cctctccttaagagc 242
|||||
Db 247 cctctccttaagagc 306
QY 243 ctatcatgtcagtlaccgctccccaacccaacactgagagagacatcttctactacagaga 302

Db	307	ctaatcatgagcaatgacgcgtccccaaccccaacaactctgagagagacatctgttactaagag	366
OY	303	tggagattagagccctctctcccaacagatctaacctgtatctctgtagccgagccaacatgtgtcg	362
Db	367	tggagattagagccctctctcccaacagatctaacctgtatctctgtagagccgagccaacatgtgtcg	426
OY	363	gattagagagcaatgtccccaagagactaacatcatgtcttcctgttgacattgagagtgctg	422
Db	427	gatacgaatgacatgttccccaagagactaacatcatctgtcttcctgttgacattgagagtgctg	486
OY	423	tcgcgtctgtgacaacgagagagaacacactctcccaacagcgcttccactgtgcaagtgctct	482
Db	487	tcgcgtctgtgacaacgagagagaacacactctcccaacagcgcttccactgtgcaagtgctct	546
OY	483	tcctcgaagctccctttgattgtgttggtgcctgagagacaatacccaagccttgtagagtaacaag	542
Db	547	tcctcgaagctccctttgattgtgttggtgcctgagagacaatacccaagccttgtagagtaacaag	606
OY	543	attactactaacccaacagagcagagctgtcccgcatgtgtttgttatcccatagacatgcatt	602
Db	607	attactactaacccaacagagcagagctgtcccgcatgtgtttgttatcccatagacatgcatt	660
OY	603	ccatcacccgcgagagaacgacctacatggtgtgcagagctgtgtctacatgataccaagagaccg	662
Db	667	ccatcacccgcgagagaacgacctacatggtgtgcagagctgtgtctacatgataccaagagaccg	726
OY	663	ctgagagatgccccttgtaaacctcccaagcgggtactagcgagatttgtatatcccttggtcttga	722
Db	727	ctgagagatgccccttgtaaacctcccaagcgggtactagcgagatttgtatatcccttggtcttga	786
OY	723	ctgccaagcgataacaacagcagacgacgtcactctctctccacaatagtgagaggtttccagct	782
Db	787	ctgccaagcgataacaacagcagacgacgtcactctctctccacaatagtgagaggtttccagct	846
OY	783	ctctgggtgagacgttatcccaagtggtcaagtttgtagggcccatggagatgttcaagatcccaagb	842
Db	847	ctctgggtgagacgttatcccaagtggtcaagtttgtagggcccatggagatgttcaagatcccaagb	906
OY	843	gtaatcagatgtatgnaaatgtgtgcatgctctcaacaaagtactatacaagaagacggttaacgtt	902
Db	907	gtaatcagatgtatgnaaatgtgtgcatgctctcaacaaagtactatacaagaagacggttaacgtt	966
OY	903	gacctatgctcaaacgtgacagcccgacgaagaagtaacggtctccgctctccctcaaacgctgcgttc	962
Db	967	gacctatgctcaaacgtgacagcccgacgaagaagtaacggtctccgctctccctcaaacgctgcgttc	1026
OY	963	acgcgctcttcgactctgtatctctgtctaacctctgagagatctcaagaagcaagactcccttcca	1022
Db	1027	acgcgctcttcgactctgtatctctgtctaacctctgagagatctcaagaagcaagactcccttcca	1086
OY	1023	ggtcatctgcgcgttgacaggttgctgtcgtcttgtagagggccctgtgtgacaactgtgatacat	1082
Db	1087	ggtcatctgcgcgttgacaggttgctgtcgtcttgtagagggccctgtgtgacaactgtgatacat	1146
OY	1083	ctctatgagcgagcgctctgggagagttgtgtatcgaactctccaaactctgcgtgagcagttcat	1142
Db	1147	ctctatgagcgagcgctctgggagagttgtgtatcgaactctccaaactctgcgtgagcagttcat	1206
OY	1143	cgatatccgcaaaccttccgtgtgcgtgcagagcgtctcggttgttgagcctggaattgtataaac	1202
Db	1207	cgatatccgcaaaccttccgtgtgcgtgcagagcgtctcggttgttgagcctggaattgtataaac	1266
OY	1203	tgaacaaggltcatgcatctgcgtctgtatgataagtcctcttgtagcccgacaactcttgaagt	1262
Db	1267	tgaacaaggltcatgcatctgcgtctgtatgataagtcctcttgtagcccgacaactcttgaagt	1326
OY	1263	gctctgcacaactctcgagatgttctcttccccgagaggcggaacattggaaccccgcaaaccc	1322
Db	1327	gctctgcacaactctcgagatgttctcttccccgagaggcggaacattggaaccccgcaaaccc	1386
OY	1323	cactgatgacgagacttttaacctcttcggccgtgtctaaatgacagtgtgaacatacaacggagt	1382

Db	1387	caagatgaacgaagacattcaaccttcggccgtgctaaatgaaacagtgaaacatcaacgaagat	1446
OY	1383	tacctctctggagatgctcgagaacggtctgtctccgcaatgtgtcccgcgacactgtgtgagat	1442
Db	1447	tacctctctggagatgctcgagaacggtctgtctccgcaatgtgtcccgcgacactgtgtgagat	1506
OY	1443	ctggcgaccttgagaacaacatccaagctgttgaaatcaacctgttcaacattcaacctgtgtga	1502
Db	1507	ctggcgaccttgagaacaacatccaagctgttgaaatcaacctgttcaacattcaacctgtgtga	1566
OY	1503	cttcgcgaatcccttctcgtcttccatctgcccgttgagtcgagaccttattgaagctgtgtct	1562
Db	1567	cttcgcgaatcccttctcgtcttccatctgcccgttgagtcgagaccttattgaagctgtgtct	1626
OY	1553	caagagatgtgtctgctgtgctctgctcgatgaagttgtctatgtgttgagggcccaactacgtcc	1622
Db	1627	caagagatgtgtctgctgtgctctgctcgatgaagttgtctatgtgttgagggcccaactacgtcc	1686
OY	1623	ttccccgtaagttctctgcctctttaccctaactgttttccatactatgctcaactctacaagt	1682
Db	1667	ttccccgtaagttctctgcctctttaccctaactgttttccatactatgctcaactctacaagt	1746
OY	1663	gggtgtctacatbgttgcaactgycacaacaactgatctcaacagagacaaagatgatgtgtct	1742
Db	1747	gggtgtctacatbgttgcaactgycacaacaactgatctcaacagagacaaagatgatgtgtct	1806
OY	1743	ttcaatgtcaactgttctcgtgtgacatagctacatacaactacaacggatltcaatgacccaatg	1802
Db	1807	ttcaatgtcaactgttctcgtgtgacatagctacatacaactacaacggatltcaatgacccaatg	1866
OY	1803	gagcctctctggagagcccccgcctctccctccctccctccctccctccctccctccctccctcc	1862
Db	1867	gagcctctctggagagcccccgcctctccctccctccctccctccctccctccctccctccctcc	1926
OY	1863	ttcagcgagcttgccatcaatgtgacccgatccagagatgtgctagtctcaacccctacgcc	1922
Db	1927	ttcagcgagcttgccatcaatgtgacccgatccagagatgtgctagtctcaacccctacgcc	1986
OY	1923	caagctgatgatgatgtgcgctgtgagagtga	1953
Db	1987	caagctgatgatgatgtgcgctgtgagagtga 2017	

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)836-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE: P129pt-F15
CLONE: P129pt-F15
US-08-232-463-14

Query Match 2.6%; Score 50; DB 1; Length 7218;
Best Local Similarity 0.5%; Pred. No. 0.00018;
Matches 2; Conservative 221; Mismatches 141; Indels 0; Gaps 0;

QY 1475 ctacacctgttcacattcaactcgttgacttcgcagatccttcgttcgctcactgcctg 1534
DB 1074 YY 1133
QY 1535 gactgcagcttataagcgtcgtcgtcctcaagatgtgtcgcgcgcgcgtcgtcgtg 1594
DB 1134 YY 1193
QY 1595 ttgtctatgttagagccactacgctcctccctccgtaagcttcgcttacttaactg 1654
DB 1194 YY 1253
QY 1655 gttctactcatgtaacatacctaagtggtgtctacatgtgtcaactgcacaaactgat 1714
DB 1254 YY 1313
QY 1715 ccacgagagccacgacatgatgctgcttcaatgctcgttcgcgtgactatgagccta 1774
DB 1314 YY 1373
QY 1775 caactacacgagctcatgaccatgagcctctcgtgagcccgccctcctcctc 1834
DB 1374 YY 1433
QY 1835 cggga 1838
DB 1434 YGTA 1437

RESULT 4
US-07-923-724-1
Sequence 1, Application US/07923724
Patent No. 5780292

GENERAL INFORMATION:

APPLICANT: Nevalainen, Helena K.M.
APPLICANT: Paloheimo, Maria T.
APPLICANT: Miettinen-Oinonen, Arja S.K.
APPLICANT: Torkkeli, Tuula K.
APPLICANT: Cantrell, Michael
APPLICANT: Piddington, Christopher S.
APPLICANT: Rambosek, John A.
APPLICANT: Turunen, Marja K.
APPLICANT: Fagerstr m, Richard B.
TITLE OF INVENTION: Production of Phylase Degrading Enzymes
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600

CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/923,724
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/496,155
FILING DATE: 19-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/044,077
FILING DATE: 29-APR-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 8610600
FILING DATE: 30-APR-1986
ATTORNEY/AGENT INFORMATION:
NAME: Cimbal, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 1050.0240004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2071 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: join(136..915, 970..1089, 1142..1245, 1305..1737)
US-07-923-724-1

Query Match 2.0%; Score 39.4; DB 1; Length 2071;
Best Local Similarity 45.6%; Pred. No. 0.12;
Matches 139; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 391 catctcgttctcgttgagacgtgagtggtgtcgcgttcgtgaaagcgagagaaacac 450
DB 444 CAGCATCAACACTACTGATATACAAAGGCGACCTGGCCTTCTGAAAGACTGGACTACTA 503
QY 451 ctctcccaacagcgttcacactgacagcgtctctctcgtcagctcccttgatgtgtgac 510
DB 504 CGTCCCTATGAGTGCTACTACACGCGGACGACCAACGAGCGCCCTACGCGGTTTGGT 563
QY 511 tgaagaaactaccagcctcgtcgtgagtgacaggaattactctaccccaagcgagctgc 570
DB 564 GAGCGGTAAACACACTGACCAAGCATTAAGGCTCGTAAGGCACTCTGTGAACGGTGA 623
QY 571 ccgcatgcttggtaacatgacatgacatgacatgacatgacatgacatgacatgacatg 630
DB 624 GACGCTGTCGCCCTTTCTTCTAGTGTGCTACGAGCTGTCTACAGAGCGCCGCAAGTT 683
QY 631 tcaagctggtgtcatatgacatgacagacccggtctgagatgctcctgaactcccaagcg 690
DB 684 CGGTGAGGTTCTTGTGCTACACTACTACCAAGCGTCCCTCAACATCATCATCGCA 743
QY 691 ctacg 695
DB 744 GTCCG 748

RESULT 5
US-08-609-426A-1
Sequence 1, Application US/08609426A

```

1 Patent No. 5830733
2 GENERAL INFORMATION:
3 APPLICANT: Nevalainen, Helena K.M.
4 APPLICANT: Paloheimo, Marja T.
5 APPLICANT: Miettinen-Oinonen, Arja S.K.
6 APPLICANT: Torkkeli, Tuula K.
7 APPLICANT: Cantrell, Michael
8 APPLICANT: Piddington, Christopher S.
9 APPLICANT: Rambosek, John A.
10 APPLICANT: Turunen, Marja K.
11 APPLICANT: Fagerstr m, Richard B.
12 APPLICANT: Houston, Christine S.
13 TITLE OF INVENTION: Production of Phytase Degrading Enzymes
14 TITLE OF INVENTION: in Trichoderma
15 NUMBER OF SEQUENCES: 69
16 CORRESPONDENCE ADDRESS:
17 ADDRESSEE: Sterne, Kessler, Goldstein & Fox
18 STREET: 1100 New York Avenue, Suite 600
19 CITY: Washington
20 STATE: D.C.
21 COUNTRY: U.S.A.
22 ZIP: 20005
23 COMPUTER READABLE FORM:
24 MEDIUM TYPE: Floppy disk
25 COMPUTER: IBM PC compatible
26 OPERATING SYSTEM: PC-DOS/MS-DOS
27 SOFTWARE: PatentIn Release #1.0, Version #1.25
28 CURRENT APPLICATION DATA:
29 APPLICATION NUMBER: US/08/609,426A
30 FILING DATE: 01-MAR-1996
31 CLASSIFICATION: 435
32 PRIOR APPLICATION DATA:
33 APPLICATION NUMBER: US 07/923,724
34 FILING DATE: 31-JUL-1992
35 PRIOR APPLICATION DATA:
36 APPLICATION NUMBER: US 07/496,155
37 FILING DATE: 19-MAR-1990
38 PRIOR APPLICATION DATA:
39 APPLICATION NUMBER: US 07/044,077
40 FILING DATE: 29-APR-1987
41 PRIOR APPLICATION DATA:
42 APPLICATION NUMBER: UK 8610600
43 FILING DATE: 30-APR-1986
44 ATTORNEY/AGENT INFORMATION:
45 NAME: Reed, Grant E.
46 REGISTRATION NUMBER: P-41,264
47 REFERENCE/DOCKET NUMBER: 1050,0080001
48 TELECOMMUNICATION INFORMATION:
49 TELEPHONE: (202) 371-2600
50 TELEFAX: (202) 371-2540
51 INFORMATION FOR SEQ ID NO: 1:
52 SEQUENCE CHARACTERISTICS:
53 LENGTH: 2071 base pairs
54 TYPE: nucleic acid
55 STRANDEDNESS: single
56 TOPOLOGY: both
57 FEATURE:
58 NAME/KEY: CDS
59 LOCATION: join(136..915, 970..1089, 1142..1245, 1305..1737)
60 US-08-609-426A-1

```

	Query Match	Best Local Similarity	Score 39.4%	DB 2	Length 2071
	Matches 139	Conservative	0	Mismatches 166	Indels 0
				Gaps	0
QY	391	catatcgttcctcgtgcacatgagatggtgtgtccgctctgtgaacacgcygagaaacac	450		
DB	444	CAGCTTCACACACTACTGTAATACAMGGGGACCTGGCCCTTCCTGAAAGATGGACCTACTCA	503		
QY	451	ctctcccaacagcgtccacattgcacgcgtctctctctcgcagctcccttgcattgattgttgagc	510		
DB	504	CTCTCCATAAGATGTCCTCTCTACACGGCCGAGACACACGGCCGCTTCAGCCGGTTTGGT	563		

Oy	511	tgaggaacataccagagccgtgcaggtataaagaattactatcccccaaggagagctgc	570
Db	564	ggaccgcgtaacacatgagcaacgatttacaaagctgcgtacagccactcttggaacggtga	623
Oy	571	ccgacatgctttagtaaccatagtaccatgcatgctcatcaccgcggagaaagcctatacgg	630
Db	624	gacggctcgccctctttttctagtggtacgacggtatcatgacagacggcccgcaagtt	683
Oy	631	tcaggtctggtgtctacatgatccagagaccggtctgagatgcccctgaacctccccagcgg	690
Db	684	cggtaggggtttcttttggtctaacactactccaccacacagctgacctctcaacatcattccga	743
Oy	691	ctacg	695
Db	744	gtccg	748

RESULT 6
 US-08-374-652C-3
 : Sequence 3, Application US/08374652C
 : Patent No. 5834286
 :
 : GENERAL INFORMATION:
 : APPLICANT: NEVALAINEN, HELENA K.M.
 : APPLICANT: PALOHEIMO, MARIA T.
 : APPLICANT: FAGERSTROM, RICHARD B.
 : APPLICANT: MIETTINEN-OINONEN, ARJA S.
 : APPLICANT: TURUNEN, MARIA K.
 : APPLICANT: RAMBOSER, JOHN A.
 : APPLICANT: PIDDINGTON, CHRISTOPHER S.
 : APPLICANT: HOUSTON, CHRISTINE S.
 : APPLICANT: CANTRELL, MICHAEL A.
 : TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
 : TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING
 : TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
 : NUMBER OF SEQUENCES: 94
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 : STREET: 1100 NEW YORK AVENUE, SUITE 600
 : CITY: WASHINGTON
 : STATE: DC
 :
 : COUNTRY: USA
 : ZIP: 20005
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/374,652C
 : FILING DATE: 24-MAY-1995
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US93/07058
 : FILING DATE: 27-JUL-1993
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/925,401
 : FILING DATE: 31-JUL-1992
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: REED, GRANT E.
 : REGISTRATION NUMBER: 41,264
 : REFERENCE/DOCKET NUMBER: 1050.071001
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 202-371-2600
 : TELEFAX: 202-371-2540
 : INFORMATION FOR SEQ ID NO: 3:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2071 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: both
 : TOPOLOGY: linear

Sequence 2 Application US/08528199
Patent No. 5763228

GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: RECOMBINANT ENZYME FOR CONVERTING
TITLE OF INVENTION: MALTOSE INTO TREHALOSE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/528,199
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,126
FILING DATE: 14-SEP-1995
APPLICATION NUMBER: US 08/485,126
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 156399/1994
FILING DATE: 16-JUN-1994

ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA=6A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-528-199-2

```

Query Match          1.8%; Score 35.2; DB 1; Length 1704;
Best Local Similarity 50.0%; Pred. NO. 1.8;
Matches      88; Conservative    0; Mismatches     86; Indels       0; Gaps        0;

OY   1760 cgytgaactatgctacaacataccagatlcatatgaccocatlgagcctctcttgaggcc 1819
Db    198 CGGGATCAGACGTCGCCGACTACACCGGGATCTCTCCGGAGATCGGCATCGAGGACTT 257
               ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY   1820 ccggcccctctccttcggagagattcgtaaatgctccggttgacttcacgcgagcttgccat 1879
Db    258 CCACGCTTCTCCTCGACGGCGGCACAGACCGCGGATCCGGGGATCATCGACTTGCTCAT 317
               ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY   1880 cacttagccgatcctcggagatgctctagctctcaacctcacgcctcgaagctgatgatg 1935
Db    318 GAACCACACGATGACGGCGCACCCCTGGTTCCAGSGCTCCCGACGATCCCGACG 373
               ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT    10
US-08-528-199-5
; Sequence 5, Application US/08528199
; Patent No. 5763228
; GENERAL INFORMATION:

```

APPLICANT: KUBOTA, Michio
 APPLICANT: TSUGAKI, Keiji
 APPLICANT: SUGIMOTO, Toshiyuki
 TITLE OF INVENTION: RECOMBINANT ENZYME FOR CONVERTING
 TITLE OF INVENTION: MALTOSE INTO TRHALOSE
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROWDY AND NEIMARK
 STREET: 419 Seventh Street, N.W., Suite 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/528,199
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/485,126
 FILING DATE: 14-SEP-1995
 APPLICATION NUMBER: US 08/485,126
 FILING DATE: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 156399/1994
 FILING DATE: 16-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BROWDY, Roger L.
 REGISTRATION NUMBER: 25,618
 REFERENCE/DOCKET NUMBER: KUBOTA=6A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-628-5197
 TELEFAX: 202-737-3528
 TELEX: 248633
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1704 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1704
 IS-08-528-199-5

[illegible]

APPLICANT: TAILOR, RH
 APPLICANT: TIPPETT, JM
 APPLICANT: BLENK, RG
 TITLE OF INVENTION: BACTERIAL GENES
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CUSHMAN DARBY & CUSHMAN
 ADDRESSEE: Intellectual Property Group of
 ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP
 STREET: 1100 New York Avenue, N.W.
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005-3918
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/286,870A
 FILING DATE: 05-AUG-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/520228
 FILING DATE: 09-MAY-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 8910624.9
 FILING DATE: 09-MAY-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: PAUL N. KOKULIS
 REGISTRATION NUMBER: 16,773
 REFERENCE/DOCKET NUMBER: 70608/220720
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 861-3000
 TELEFAX: (202) 822-0944
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2159 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:

```

US-08-286-870A-7
1:11:1029

Query Match      1.8% ; Score 34.8; DB 3; Length 2159;
Best Local Similarity 49.5%; Pred. No. 2.7;
Matches 90; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 592 ccaatgcacatgcatcaccagccgcagaagacgcgtctacattggtcaagctgtgtctacatgat 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1170 CCTCCTCCCTTCACGACCGCGAGCGTGTACCAGCACGAGACCGCTGGCGGCCTGAACCTGTT 1229

QY 652 ccaagaccggcgctgagtatgccctgaacctcccaccaggcgctacagtcgagtttgatatccc 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1230 CCTGATACCAAGCCCCGTGAACGGCGGTGCCCCC GGCTGGACTTTTCACTTGAAAGTTCGTGACCCA 1289

QY 712 ctgtgtctgactcctcacaagcagatacaagcacagacgcgcactctctctccaaactlgyaga 771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1290 CCGCATGGCGACGAGCAACAATTCTTA CTNACC CGCGGTACGCTGSCATTGGCACCCAACTCCA 1349

QY 772 gg 773
    ||
Db 1350 GG 1351

RESULT 13
US-08-446-922-5/C
Sequence 5, Application US/08446922
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 20:23:09 ; Search time 4997.71 Seconds
(without alignments)
8772.237 Million cell updates/sec

Title: US-09-656-640A-3
Perfect score: 2095
Sequence: 1 cagctcggtctactctc.....gaataagttgtgtgcttaa 2095

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pal:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution:

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

1	2095	100.0	2095	6	ARI22997	ARI22997 Sequence
2	2095	100.0	2095	6	AX101098	AX101098 Sequence
3	2095	100.0	2095	6	AX108674	AX108674 Sequence
4	1946.2	92.9	1958	6	ARI22996	ARI22996 Sequence
5	1946.2	92.9	1958	6	AX101096	AX101096 Sequence
6	1946.2	92.9	1958	6	AX108672	AX108672 Sequence
7	528.6	25.2	1942	8	MYRBR	D12579 M. verrucar
8	528.6	25.2	1959	8	E05283	E05283 Biliirubin o
9	517.4	24.7	2126	8	AMU271104	AJ271104 Acromoni
10	507.6	24.2	3241	8	MYRBD	D14081 M. verrucar
11	95.8	4.6	1422	6	AX195966	AX195966 Sequence
12	95.8	4.6	109519	6	AX195929	AX195929 Sequence
13	82.8	4.0	3446	6	AX204969	AX204969 Sequence
14	81.4	3.9	9531	1	AY033994	AY033994 Streptomy
15	79.8	3.8	137879	8	AP003233	AP003233 Oryza sat
16	79.8	3.8	141528	8	AP002909	AP002909 Oryza sat
17	73.2	3.5	22302	1	SAPRHA	U04283 Streptomyce
18	69.8	3.3	134673	8	AP002860	AP002860 Oryza sat
19	61.2	2.9	9413	8	LDMOFA	225774 Lepidothrix
20	57.8	2.8	134673	8	AP002860	AP002860 Oryza sat
21	51.6	2.5	141528	8	AP002909	AP002909 Oryza sat
22	50	2.4	7218	6	I66494	I66494 Sequence 14
23	49.6	2.4	1533	6	AX121137	AX121137 Sequence
24	49.6	2.4	1614	6	AX065821	AX065821 Sequence
25	49.6	2.4	349980	6	AX127146	AX127146 Sequence
26	48.2	2.3	125020	9	AF429315	AF429315 Homo sapi
27	47.2	2.3	3170	10	RNNCAM14	X06564 Rat mRNA fo
28	46.4	2.2	15897	1	BSU51115	U51115 Bacillus su
29	46.4	2.2	16585	1	AB007638	AB007638 Bacillus
30	46.4	2.2	213190	1	BSOB0004	Z99107 Bacillus su
31	45.6	2.2	2544	10	BC011310	BC011310 Mus muscu
32	45.6	2.2	3574	4	BTADCYC	X16451 Bovine mRNA
33	45	2.1	19835	2	AC092489	AC092489 Oryza sat
34	44.8	2.1	10732	6	E32986	E32986 Gene encodi
35	44.6	2.1	10738	1	AE009694	AE009694 Brucella
36	44	2.1	2583	10	MNMCAMR	X15049 Mouse commo
37	44	2.1	2583	10	MNMCAMR	Y00051 Mouse mRNA
38	43.6	2.1	127485	9	HUAC002039	AC002039 Homo sapi
39	42.6	2.0	125020	9	AF429315	AF429315 Homo sapi
40	42.4	2.0	1882	10	NM028495	U28495 Mus musculu
41	42.4	2.0	2031	10	NMRHOBIN	X95761 M. musculus
42	42.4	2.0	3131	10	AF177032	AF177032 Mus muscu
43	42.4	2.0	3154	10	BC006589	BC006589 Mus muscu
44	42	2.0	129473	2	AC040158	AC040158 Homo sapi
45	42	2.0	143900	2	AC092137	AC092137 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS ARI22997 2095 bp DNA
DEFINITION Sequence 3 from patent US 6168936.
ACCESSION ARI22997
VERSION ARI22997.1 GI:14107963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2095)
AUTHORS Wang,H.
TITLE Phenol oxidizing enzymes
JOURNAL Patent: US 6168936-A 3 02-JAN-2001;
FEATURES
source location/Qualifiers
1..2095
/organism="unknown"

BASE COUNT 437 a 618 c 510 g 530 t
ORIGIN

Query Match	Score	2095;	DB 6;	Length	2095;
Best Local Similarity	100.0%;	Pred. No. 0;			

DEFINITION Sequence 3 from Patent WO0121809.
ACCESSION AX101098
VERSION AX101098.1 GI:13619952
KEYWORDS
SOURCE Stachybotrys chartarum.
ORGANISM Stachybotrys chartarum.
REFERENCE 1 (bases 1 to 2095)
AUTHORS Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Stachybotrys.
TITLE Wang, H.
JOURNAL Stachybotrys phenol oxidizing enzyme
PATENT: WO 0121809-A 3 29-MAR-2001.
GENBANK INTERNATIONAL, INC. (US)
FEATURES
source
Location/Qualifiers
1..2095
/organism="Stachybotrys chartarum"
BASE COUNT 437 a 618 c 510 g 530 t
ORIGIN

Query Match 100.0%; Score 2095; DB 6; Length 2095;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2095; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagctcgctactactctcgtctctctcttgacaatacattacaaatcgtctctca 60
DB 1 CAGCTCGGCTACTACTCTCGCTCTCTTGACAAATCAATCTACCAATCGTCTTCA 60
QY 61 attcaacgatcaacaatgatacgaacgatacggagccgtggtctcgtggtcgt 120
DB 61 ATTTCAAAGATCAACATATCAGCCAACTACGAGCGGTGCTGCTGCTGCTGCT 120
QY 121 gcttcgagcgagctcgtctgatacgaatccgtctgctgctgatacgaatcgtctc 180
DB 121 GATTCGGCGGACGCTCTGCTGATGCCAGATCCGTTGCTGCTGCTGCTGCTGCT 180
QY 181 cgtctcacaagaagagcagacgagctgagctccctcgtcgtctgatacgaatcgtcct 240
DB 181 CGGTCTCACCAAGAGCAGACGAGCTGAGTCCCTCCCTGCTGCTGCTGCTGCTGCT 240
QY 241 gccgatacctctctgaaagcgcccaagtgaatgatacttataagctagacagca 300
DB 241 GCCGATCCCTCTCTGAAAGCGCCCAAGTAAATGATCTTATGCTGAGCAGAGCA 300
QY 301 acgttgataatcagtcagctacccgtcccaacccaacccaacccaacccaacccaac 360
DB 301 ACCTTGCTAATCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 360
QY 361 acgagatgagataagagccctctccaccagatacctacccgtgatacggagcgagca 420
DB 361 ACGAGATGAGATGAGGCTCTCTCCACAGATCACTGATCTGAGAGCGGCGCAACA 420
QY 421 tggcttgagatacgaatgagctgctcccaagatacctacccgtcgtcgtgagcaatg 480
DB 421 TGGTTGGATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 ggtgtgtcgcgtctgtaagacgagagaaacacccctcccaacgagcgtccacgtcag 540
DB 481 GGTGTGTCCGCTTCGTAAGACGAGAGAAACCTCTCCCAACAGCTGCTCAATTCACAG 540
QY 541 gctctctctcagagctccctctgagctggtggtgagagacatacgaacccgtgagat 600
DB 541 GCTCTCTCTCTGAGCTCCCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 acaagagatacctaatacacaagaagcgtgcccgaatgcttggtaacatgacatg 660
DB 601 ACAAGGATTAACATACCCCAACAGGAGCTGCCGATGCTTTGATGATGATGATGATGAT 660
QY 661 ccatgtccatcagccgagagaaacgctacatggtgacagctgtgtctacatgatacag 720
DB 661 CCAATGCTCAACCCGAG 720
QY 721 acccgagctgagagatgacctcccaacccagcggtacagcgagttgatatcccttgg 780

DB 721 ACCGGCTGAGAGATGCGCTTAACCTCCCGAGCGCTAACGCGAGTTGATATCCCTTGG 780
QY 781 ttctgagtcgcaagcagatacaacgagcagcagcagcagcagcagcagcagcagcagc 840
DB 781 TTCTGAGTCGCAAGCAGATTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 ccagctctcggagtgagcagctatcaagtggttaagttgagccattgagatgctcagatc 900
DB 841 CCAAGCTTCGGGTGAGAGCTTATTAAGTGTAGTGAAGCCCAATTGAGATGCTTCAGATC 900
QY 901 ctagaagatcgaatgataagaatgagatgagatgagatgagatgagatgagatgagatg 960
DB 901 CTAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 gctctgagctatgatacagtgagcagcagcagcagcagcagcagcagcagcagcagc 1020
DB 961 GCTTGGCTATGCTCAACCTGAGCGCGCAAGTACCGCTTCGCTTCCTCAACGCTGC 1020
QY 1021 cgtctcagcgtctcgtctgatacttgctacattgagagatcagaagcagcagcttc 1080
DB 1021 CGTCTACGCTCTTTCGCTCTGATCTGCTACCTCTGAGGATTCAGAGACGAGCTTCC 1080
QY 1081 ctccagagctatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
DB 1081 CTTCAGAGCTATGCGCGTSGAGGTGTGCTGAGAGGCGCTGTTGACACTGACACTCT 1140
QY 1141 gtaacatctatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200
DB 1141 GTACATCTCTATGCGCGAGCGAGCTGAGAGGTTGATTCAGCTTCCACCTTCCTGCGCA 1200
QY 1201 gtcacatgatacagcaacctctcgtctgagcagcagcagcagcagcagcagcagcagc 1260
DB 1201 GTCCATCGATATCCGACACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 taacactgacaagagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320
DB 1261 TAACACTGACAAGCTATGAGGATTCGCTGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 tgaagtgctcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380
DB 1321 TGAAGTGCTTCGCGACCTCCGAGATGTTCTTCCCGAGGCGGCAACCTGCGACCCCG 1380
QY 1381 aaacccacatgatacgaagacatcctcagcagcagcagcagcagcagcagcagcagc 1440
DB 1381 AAACCCACTGATGAGAGAGACTTTCACCTTGGCCCTGCTAAATGAGACAGTGACAATCA 1440
QY 1441 cgaagttacatcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1500
DB 1441 CGAGATTACCTTCGAGATGTCAGAGAACCGTCTGCTCCGAAATGTGCCCGCACACTGT 1500
QY 1501 tgaagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560
DB 1501 TGAGATCTGCGAGCTTGAGAGAACTTCAACAGGTTGAGACTCACCTGTTCACTTACCT 1560
QY 1561 cgttgatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1620
DB 1561 CGTTGACTTCGAGCTTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1621 tggctcagaagatgctgctgagcagcagcagcagcagcagcagcagcagcagcagcagc 1680
DB 1621 TGGTCTCAAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
QY 1681 cgtctctcccgtaagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1740
DB 1681 CGCTCTTTCCTGAGTGTCTGCTGCTTAACTTAACATGATGATGATGATGATGATGAT 1740
QY 1741 acaagtgagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800
DB 1741 ACAAGTGAGTCTACATGTTGACCTGCAACCTGATCCACAGGACCAAGCAATGATG 1800
QY 1801 gctgcttcaatgacatctcgtcgtgatacgtgatacagcagcagcagcagcagcagcagc 1860

|||||
Db 1203 TGACAAAGCATGATGGATTCCTGTTGATGAAGTCCTTGAGTCGCCCAACACTTTTGAGCT 1262
Qy 1327 gcttgcgaacccctccgaagatcttccctccgaagagcggaacatggggaacccggaaccc 1386
Db 1263 GCCTGCCAACCTCCGAATGTTCTTCCCGAGGGGGGACACTGGGACCCCGGAACCC 1322
Qy 1387 cactgatacgaagacttccactccgctgctgaatgagatggaacatcaacgaagat 1446
Db 1323 CACGATGACGACACTTTCACCTTCGGCCGGGTAAATGAGACAGTGAACAAACGAGCT 1382
Qy 1447 tacccttcgagatgcgaagaacgctgctccggaatgctcccggaacactgctgagat 1506
Db 1383 TACCTTCTCGATGTCGAGAACCTGTGCTCCGCAATGTCGCCGACACTGTTGAGAT 1442
Qy 1507 ctgagcacttgaagaacaccccaacaggttgagacccactgttcaacttccactgttga 1566
Db 1443 CTGGCGACTTGAGAACAACTCCAAACGGTGGACACCCCTTTCACATTCACCTGTTGA 1502
Qy 1567 ctccgaagcttccctcgttccactgcccgtgagagtcgaagccttaagagctgctgct 1626
Db 1503 CTTCGAGATCCTTCTGCTGCTCACTGCCCGGAGTGGAGCTTATGAGGCTGCTGCT 1562
Qy 1627 caaggaatgcttccgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1686
Db 1563 CAAGGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1622
Qy 1687 ttcccgtaagcttccgcttccactaacgctgttccactgacttcaacttcaacttcaagt 1746
Db 1623 TTTCCCGTAATGCTTCCGCTTTTACTACTAGTGTTCCTTACCTATGCTTAACATCTAAGT 1682
Qy 1747 ggtgttcaatgcttgcactgccaacactgaatccaagagagacacagatgagctgct 1806
Db 1683 GGTCGTACATGTTGACACTGTCACAACTGATCCAGAGGACCAAGCATATGCTCTCT 1742
Qy 1807 ttcaatgctatgcttccgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1866
Db 1743 TTCAATGCTACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1802
Qy 1867 gaggctctcgaagagcccgcccttccctcctcctcgaagagcttcaagaatgctcgaag 1926
Db 1803 GAGCTCTCTGAGAGCCCGCCGCCCTCTCTCTCGAGAGATTGGAATAGCGCTGAGC 1862
Qy 1927 ttcaagagagcttgcatacactgaacccaatccaagagatgctgacttcaacccctgaagc 1986
Db 1863 TTCAAGCAGCTTGGCATCATCAGCAGCATTCAGGAGATGGCTTCAACCCCTTACGCC 1922
Qy 1987 caggctgatagatacgcctgagagagtaaa 2017
Db 1923 CAGGCTGATGATGATGCCGCTGAGGAGTGA 1953

RESULT 5
AX101096 1958 bp DNA Linear PAT 10-APR-2001
LOCUS AX101096 Sequence 1 from Patent W00121809.
DEFINITION AX101096
ACCESSION AX101096
VERSION AX101096.1 GI:13619951
KEYWORDS
SOURCE Stachybotrys chartarum.
ORGANISM Stachybotrys chartarum.
Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Stachybotrys.
REFERENCE 1 (bases 1 to 1958)
AUTHORS Wang, H.
TITLE Stachybotrys phenol oxidizing enzyme
JOURNAL Patent: WO 0121809-A 1 29-MAR-2001;
GENENCOR INTERNATIONAL, INC. (US)
FEATURES
source Location/Ouallifiers
1..1958
/organism="Stachybotrys chartarum"
/db_xref="taxon:74722"

BASE COUNT 394 a 593 c 490 g 481 t
ORIGIN

Query Match 92.9% Score 1946.2; DB 6; Length 1958;
Best Local Similarity 99.8% Pred. No. 0;
Matches 1948; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 67 aagcatacaatgatacgaacagatcagagccgtgctcgtgagcttctgagatcag 126
Db 3 ATCCATCAACATGATTCAGCCCAAGCTATCGAGCGCTGCTGGCCCTGTGTGATCGG 62
Qy 127 cggagctctgtcgtatgcagatccgttctgtgcatcgaacagatgcttccgctc 186
Db 63 CGGAGCTCTGTGCATGAGCATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 122
Qy 187 caccagaagcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 246
Db 123 CACCAAGAGGAG 182
Qy 247 cccctccttgaagcggcccaagtagtaagatcattcttaagctagcagagccaagctg 306
Db 183 CCTTCCTTGAGGCGCCCAAGTAGTATGATTTCTATAGCTAGAGCAACCAAGCTTG 242
Qy 307 ctatcactgagtaacgcttcccaaccccaacactgagagagacatctgtactaagaga 366
Db 243 CTAAATCATGTCAGTACCGTCCCAACCCCAACACTGAGAGAGATCTTGTACTACGAGA 302
Qy 367 tggagattagagcccttcccaacagatcactcctgactcgtgagcggccaacatgttg 426
Db 303 TGGAGATTAGGCCCTTCTCCCAAGATCTACCTGATCTGAGCGGCGCAACATGTTG 362
Qy 427 gatacagatgagatgctcccaagagactacatcatcgttccctcgttgcactgagagtg 486
Db 363 GATACGATGAGCATGCTCCCAAGACCTACCATCATGCTTCTCTGCTGCTGCTGCTGCT 422
Qy 487 tccgctcgtgaacagcggagagaaacacttcccaacagcgttccacttgaagcctt 546
Db 423 TCCGCTTGTGAAGCGGAGAGAACACTCTCCCAACAGGCTCACTGACAGGCTCTT 482
Qy 547 tctctgagcttcccttgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 606
Db 483 TCTCTGAGCTCCCTTTTATGTTGGGTGAGAGAACATACCAAGCTGCGAGTCAAGG 542
Qy 607 attactactacccaacagagcagcgttcccgatgcttctgtgaacatgacatgcatgt 666
Db 543 ATTACTACTACCCCAACAGGACGCTGCCGATGCTTTGTGATCATGACATGCGCATGT 602
Qy 667 ccatcaccgcccgaagaaacgactacatcaggtcaggtgtgtgtgtgtgtgtgtgtgtgt 726
Db 603 CCATCACCAGCGAGAAACGCTCATGAGGTCAAGGCTGCTGTACTGATCCAGGACCCGG 662
Qy 727 ctgagagatgaccttgaacctcccaacgctcagcgagatgttgaatctccctgtgttga 786
Db 663 CTGAGGATGCTCTGAACTCCCAAGCGCTACGGGAGGAGTTGATATCCCTTGTGTTGTA 722
Qy 787 ctgcaagcagatatacagaacagcagcactcttctcccaacaaatgagagtgctcagct 846
Db 723 CTGGCAAGGATACAAACGAGAGGACACTCTTCTTCCACCAATGAGAGGTTTCCAGCT 782
Qy 847 tctggggtgagcttactcaagtgtgaagtgtgagcccatgtgagatgtcttaagatct 906
Db 783 TCTGGGGTGACGCTTATTCAGTGAAGTGAAGCCCATTTGAATGATCTTCAGTACGAA 842
Qy 907 gtaatgataatgaatgtgtgactgtcctaaccagatgtcattacagaagaagctgacgct 966
Db 843 GATACGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902
Qy 967 gctatgctcaagctgcaacgagcagcagcagcagcagcagcagcagcagcagcagcagc 1026
Db 903 GCCTATGCTCAACGTCGACCGCGCAAGTACCGCTTCCGCTTCTCAACGTCGCGCTTC 962
Qy 1027 acgcttcttgcctgttacttctgttactccttgaagatltcaagagacagacttcccttca 1086
Db 963 ACGCTCTTCCGCTGTATCTGTACTGATCTGAGATTCAGAGACGAGACTTCCCTTCCA 1022


```

QY 1087 ggtcattgcccgtgacggtggtctgtctttagggccctgtgacactgacactctgtacat 1146
    |||||||
Db 1023 ggtcattgcccgtgacggtggtctgtctttagggccctgtgacactgacactctgtacat 1082
QY 1147 cctctatgcccagcgcgtgaggaggtgtatcgaactctccactctgctgagcagtcacat 1206
    |||||||
Db 1083 ctctatgagccgagccctggagaggtgtatcgaactctccactctgctgagcagtcacat 1142
QY 1207 cgaattccgcaacactctcctggtgctgacggtgtcgtgtgtgtgagctgagttgataaac 1266
    |||||||
Db 1143 cgaattccgcaacactctcctggtgctgacggtgtcgtgtgtgtgagctgagttgataaac 1202
QY 1267 tgacaagatgacgacatcgtctgtatgaagaatccctgtgctgacgacactctgagat 1326
    |||||||
Db 1203 tgacaagatgacgacatcgtctgtatgaagaatccctgtgctgacgacactctgagat 1262
QY 1327 gcttgccaaacctcgaagatgtctcttcccgaggcggaactgagaccccgcaaaccc 1386
    |||||||
Db 1263 gcttgccaaacctcgaagatgtctcttcccgaggcggaactgagaccccgcaaaccc 1322
QY 1387 cactgatacagagactcttcaactctgcccgtgtctatgagacagtgagacaatacagagat 1446
    |||||||
Db 1323 cactgatacagagactcttcaactctgcccgtgtctatgagacagtgagacaatacagagat 1382
QY 1447 tacctctcgcgagatgacgacacgctcgtcgcgaatgtgcccgcgacactgttgaat 1506
    |||||||
Db 1383 tacctctcgcgagatgacgacacgctcgtcgcgaatgtgcccgcgacactgttgaat 1442
QY 1507 ctgagcgaacttgaagaacactccaaagttggaactaacctgttcaactaacactcgttga 1566
    |||||||
Db 1443 ctgagcgaacttgaagaacactccaaagttggaactaacctgttcaactaacactcgttga 1502
QY 1567 ctccgaagctccttctgtctgacactgcccgtgagtgagtcgaagccttatgagcgcgtgtct 1626
    |||||||
Db 1503 ctccgaagctccttctgtctgacactgcccgtgagtgagtcgaagccttatgagcgcgtgtct 1562
QY 1627 caaagatgtgtcgtgagcgtcgtcgtcgtgagatgtgtcctatgttgaagcccactacgtcc 1686
    |||||||
Db 1563 caaagatgtgtcgtgagcgtcgtcgtcgtgagatgtgtcctatgttgaagcccactacgtcc 1622
QY 1687 ttcccgtaagttctcgtcgtccttaccactaactggtttcaactaactaactaactaact 1746
    |||||||
Db 1623 ttcccgtaagttctcgtcgtccttaccactaactggtttcaactaactaactaactaact 1682
QY 1747 ggtgtctacatgtttgacactgcccacaaactgatacgaagagacacgaactgtatgtctgt 1806
    |||||||
Db 1683 ggtgtctacatgtttgacactgcccacaaactgatacgaagagacacgaactgtatgtctgt 1742
QY 1807 ttcaatgtcaactgttctcgtgtgactatggtctaacactaacagagttcaatgacccatg 1866
    |||||||
Db 1743 ttcaatgtcaactgttctcgtgtgactatggtctaacactaacagagttcaatgacccatg 1802
QY 1867 gaggcctctctgagagcccccgcctctcctcctcgtgagagttcgaagaatggtcgtgagac 1926
    |||||||
Db 1803 gaggcctctctgagagcccccgcctctcctcctcgtgagagttcgaagaatggtcgtgagac 1862
QY 1927 ttcaagagagctgtgacatcaactgaacgcgacattcaagagagatggttagcttcaacccctaa 1986
    |||||||
Db 1863 ttcaagagagctgtgacatcaactgaacgcgacattcaagagagatggttagcttcaacccctaa 1922
QY 1987 caggctgagatgagatgacgagcgtctgagagatga 2017
    |||||||
Db 1923 caggctgagatgagatgacgagcgtctgagagatga 1953

```

```

RESULT 6
AX108672 1958 bp DNA linear PAT -30-APR-2001
LOCUS AX108672
DEFINITION Sequence 1 from Patent WO0121748.
ACCESSION AX108672
VERSION AX108672.1 GI:13923900
KEYWORDS

```

```

SOURCE Stachybotrys chartarum.
ORGANISM Stachybotrys chartarum.
REFERENCE 1 (bases 1 to 1958)
AUTHORS Convents,D.U., Doornink,M.O., de Vries,C.H. and Wang,H.
TITLE Detergent compositions comprising phenol oxidizing enzymes
JOURNAL Patent: WO 0121748-A 1 29-MAR-2001;
FEATURES
    source location/Qualifiers
        1..1958
        /organism="Stachybotrys chartarum"
        /db_xref="taxon:74722"
BASE COUNT 394 a 593 c 490 g 481 t
ORIGIN
Query Match 92.9%; Score 1946.2; DB 6; Length 1958;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1948; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 67 aacgatacaatgatgacgcaagatcagagccgtgagctctgagccttgatagc 126
    |||||||
Db 3 atccatcaacatgatgacgcaagatcagagccgtgagccttgatagc 62
QY 127 cggcagctctgcatgacagatccgtgtgctgcatgacagacatgctccgtct 186
    |||||||
Db 63 cggcagctctgcatgacagatccgtgtgctgcatgacagacatgctccgtct 122
QY 187 caaccaagagcagcagcagctgagctccctcctgcttgaacgaatgctctgcgat 246
    |||||||
Db 123 caaccaagagcagcagcagctgagctccctcctgcttgaacgaatgctctgcgat 182
QY 247 cctctctgaaagcggcccaagtaagtaagtaagtaagtaagtaagtaagtaagtaag 306
    |||||||
Db 183 cctctctgaaagcggcccaagtaagtaagtaagtaagtaagtaagtaagtaagtaag 242
QY 307 ctaatcaatgacagacagcagcagcagcagcagcagcagcagcagcagcagcagc 366
    |||||||
Db 243 ctaatcaatgacagacagcagcagcagcagcagcagcagcagcagcagcagcagc 302
QY 367 tggagatgagccctctctccacacagatctacccgtatctgagccgcgaatggttg 426
    |||||||
Db 303 tggagatgagccctctctccacacagatctacccgtatctgagccgcgaatggttg 362
QY 427 gatagatgagatgagcccaagacacacacacacacacacacacacacacacacacac 486
    |||||||
Db 363 gatagatgagatgagcccaagacacacacacacacacacacacacacacacacacac 422
QY 487 tccgctctgtagacagcggagagaacactctcccaacagcgtcccaatgacagcgtct 546
    |||||||
Db 423 tccgctctgtagacagcggagagaacactctcccaacagcgtcccaatgacagcgtct 482
QY 547 tctctcagactcctcttgatggtgtgagtgagacactcagcagcctgagcgtgacaa 606
    |||||||
Db 483 tctctcagactcctcttgatggtgtgagtgagacactcagcagcctgagcgtgacaa 542
QY 607 attactacaccccaacagcagcgtgagcagcagcagcagcagcagcagcagcagcagc 666
    |||||||
Db 543 attactacaccccaacagcagcgtgagcagcagcagcagcagcagcagcagcagcagc 602
QY 667 ccataccgcgcgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 726
    |||||||
Db 603 ccataccgcgcgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 662
QY 727 ctgagatgagccttgaaactcaccagcagcagcagcagcagcagcagcagcagcagcagc 786
    |||||||
Db 663 ctgagatgagccttgaaactcaccagcagcagcagcagcagcagcagcagcagcagcagc 722
QY 787 ctgcaagagatatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 846
    |||||||
Db 723 ctgcaagagatatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 782
QY 847 tctggggtgagcttattcaagtgtgagtgagccatgagatgattagatcctagaa 906
    |||||||

```

Db	783	TCGCGGAGTGCAGCTTATTCGAAGTGAAGTTGAGCCCATTTGAGATGCTWCAAGATCTCGA	842
Oy	907	gtatcgatgatafgaaattgttgatgctctaaacagtgatatacaagaaggttaagccttg	966
Db	843	GATTCAGATGATGAATTAATTTGATGCTCTTACCAATGCTATACAGAGCGGTACGCTTG	902
Oy	967	gactatgctcaacgctcaagccgcgaagtaacgccttcgccttcctcaacgctgacgtctc	1026
Db	903	GCGTATGCTACAGTGCACCGCGCAAGTACCGCTTCGGCTTCCTCAACGCTGCCGCTCC	982
Oy	1027	acgcctcttcgcctctgatactgtactctgaagattcaagaatcagaagacgaacttcctcca	1086
Db	963	ACGCTCTTTCGCTCTTATCTTCTGTAACCTGATGAGATTCAGAGACAACACTTCCTTTCA	1022
Oy	1087	ggtcatatgcgcctgaaggtgagctcctctgaaggccctgtgaacatgaacatctgatac	1146
Db	1023	GGTTCATTGCCGCTGACGGGTGCTGCTTGAAGGCCCTGTTGACACTGACACTCTGTATAT	1082
Oy	1147	ctctatgcccgaagcgcctgggaagtgatgtatacgaattctcaactcgaacttcgcctgac	1206
Db	1083	CTGTATGGCCGACGCTGGAGAGTTGTATTCAGACTTTCACACTTCGCTGCCAGTCCAT	1142
Oy	1207	cgaataccgaacacttcctcgtggtgacaggtctcgtgtttgaagctgaagttgaatacac	1266
Db	1143	CGATATCCGAACCTTCCTGCTGCTGCTACGGCTCTCGAGTGTAGGCTTGAGCTTTGATAAC	1202
Oy	1267	tgacaagtcatacgcatcgtcgtcgttgatgaagtcctctgaagtcgcgcgaacattcgaag	1326
Db	1203	TGACAAGTCAATGCGATGCTGTGTTGATGAAAGTCCCTGATGAGTGCGCCGACACTTCGAG	1262
Oy	1327	gcctgcgaacctccgaagatgcttccttcccccgaaggcgaactbgagcccccgaaaccc	1386
Db	1263	GCCCTCCCAACCTCCGAGATGTTCTTTTCCCGAGGCGGCAACTGGAGCCCGCAAAACC	1322
Oy	1387	cactgaatgaagaaacttcaacttcgcgcctgtctaatagacagtgagaatcaacaggaat	1446
Db	1323	CACGATGAGAGAACTTTACCTTGGGCGGTCTCATGAGACAGTGGACAATCAACGGAGT	1382
Oy	1447	taaccttcgaatgacgaagaacacgctcgtctccgcaatgtgcccgcgaacatggttgaat	1506
Db	1383	TACCTTTCGAGATGCGAGAACCGTCTGCTCCGCAATGAGTGCACCACTGTTGAGAT	1442
Oy	1507	ctggcgaacttgaaagaacacccaagcttgagactcaaccctgttcaatacctccgttga	1566
Db	1443	CTGGCGACTTGAGAACACTCCACACGGTTGAGACTACCTGTTCAATTCACTCCTGTTGA	1502
Oy	1567	cttccgaagtccttctcgtctcaactgcgcctgaggtcgagccctatgagctgctgctct	1626
Db	1503	CTTCCGAGTCTTCTTCTGTTCCACTCCCGGTGAGCGAGCTTATGAGGCTGCTGCTCT	1562
Oy	1627	caagaatggtgtcgtcgtgctctgtctgtgaagtttctatgfttgaagcccaactaagctcc	1686
Db	1563	CAGAAGATGTTGCTGCTGCTGCTGCTGCTGAGGTGTATATGTTGAGGCCACACTACGCTCC	1622
Oy	1687	tttcccgtaagttctcgccttcttacctaaactctggttttcaactatgcttacaactcaaat	1746
Db	1623	TTTCCCGAATGTTCTGCCCTTTTACCACTAGGATTTTCACTCATGCTAACACTCTCAAGT	1682
Oy	1747	ggtgtctacaatgtgtgcaactgcacaacacgtatccaagaagaccagaacatgatatgctct	1806
Db	1683	GGTGTCTACATGTTGCACTCCCAACCTCATATCCAGAGACACGACACTGTATGCTCTCT	1742
Oy	1807	ttcaatgtcaactgtctcgcgttgaaactggttacaactacaacgaagtttcatltgaacccatg	1866
Db	1743	TTTCATATGCACTGTTCTCGGTGACTATGAGTGAACATTACACCGAAGTTCAATTACCCATG	1802
Oy	1867	gaagcctctctgagagcccgccctctcctcctcgcgaagttcgaagaatgctcgggtgac	1926
Db	1803	GAGCTCTCTGGAAGCCGCCCTTCTCTCTCCGAGAGTTGAGAGATAGCTCGGGGTAC	1862
Oy	1927	ttcgaagagcttgcatacaactgaacgcatccaagaatgctcgaacttcaaccctcgaagc	1986
Db	1863	TTTCAACGAGTCTGCATCTACGACCGCATTCAGAGAAATGCTATGCTTCAACCCCTTACGCC	1922

[illegible]

Db 1334 GCTGACTTATGGATGAAAGACGTATACCAATCTGTGAATGAGACAAGATCTGTACT 1393
OY 361 aagagatgagatgagcctctccaccaatctacactatctcgagcgccgaaca 420
Db 1394 ATGAGGTGAGATCAAGCCCTTACACACAGGTTTACCTGACCTGGATCCGCTGATC 1453
OY 421 tggatgatacagatgcatgctcccaagacatacaactgctccctggagc----- 473
Db 1454 TGGTGGGATATGAAATGTCGCGCCCTACTTCCAGGTCTCTGGTGAACATT 1519
OY 476 -----tgagag 481
Db 1514 CGAGACCATGCAAAAAGCATTGAAATGACGCTGATTAGAAAACAGATGTTGAAAC 1573
OY 482 tggatgagatgagcctctccaccaatctacactatctcgagcgccgaaca 541
Db 1574 ATGTGTCGCTTATTAACAAAGCTGAGG-----CTCTTAACCTCGTTCCCTGCACGG 1627
OY 542 cctctctcagagcctccttgatggtggtgagagacatacccaagcctgagcagta 601
Db 1628 ATCATTTCTCTGTCGCGCTTTGACGAGATGGCGAGAGACATCCGACCTGGCAGCTT 1687
OY 602 caagattactactacccaacagagcgtgcccagatgcttggatgatacagcagtc 661
Db 1688 CAAAGACTTATCAACCAATAGACAGCTGCTGCTACCTATGATGATACAGATCATGC 1747
OY 662 catgcatcacaacgagagagcagcagcagtgagtgagtgatggtgataatcagta 721
Db 1748 TATGATATACCTGCTGAGAACGCTACCGCTGCGCTGCTGTATATGCTCTCACTGA 1807
OY 722 cccgagctgagatgacctcccaagcagcagcagcagtgatgatacctccttgg 781
Db 1808 CCAGCGCAAGAGCGCTCACTTGCAGATGGATATGGCGAGTTGCATATCCATATGAT 1867
OY 782 tctgactgcgaagcagatatacagcagagcagcctctctccacaatgagagtttc 841
Db 1868 CCTCAGCTGCAAGCAATATACCGCAAAAGCACTGTGACACACTAATGAGAGGTGAA 1927
OY 842 cagctctgagtgagcagcagcagcagcagtgatgagtgagtgagtgagtgagtc 901
Db 1928 CTGATTTCTGGGATGATGATATACCTGCTGATGATGATGATGATGATGATGATG 1987
OY 902 tagaagatcagatgataatgtagatgtagatgtagatgtagatgtagatgtagatg 961
Db 1988 ATACATATACACCCCATTAACAGTACGATATGACATGACAACTGT---AGAAAGTCA 2044
OY 962 ccttgagcagatgtagcagcagcagcagcagcagcagcagcagcagcagcagc 1021
Db 2045 CCTGGCCTTTGAGAACGCTGAGCCTCGCAATATGATGATGATGATGATGATGATG 2104
OY 1022 gftcagagcctctgagcagcagcagcagcagcagcagcagcagcagcagcagc 1081
Db 2105 GTTCTCTGCTTCTTGGCTTACTTGTGATGATGATGATGATGATGATGATGATG 2164
OY 1082 ttcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1141
Db 2165 TTCAAGGTTATGCGCTGCTGCTTCTTGAACACCGTCGATACGAGCTTGCAG 2224
OY 1142 tacaatcctagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1201
Db 2225 TACATTTCCATGCGAGGAGGTTGATGATGATGATGATGATGATGATGATGATG 2284
OY 1202 tccatagatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1258
Db 2285 ACTATTGAACTCCGCACTGGGCGGTACATTTGGCGCATGGAACAGATACCGCTAT 2344
OY 1259 gataacatgacaagatgtagatgtagatgtagatgtagatgtagatgtagatgtag 1318
Db 2345 GACAACACGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2404
OY 1319 tctgagtgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1378

Db 2405 TCAGTTGTCCTGCTAACCTTGATGATGATGATGATGATGATGATGATGATGATG 2464
OY 1379 gaaacccacagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1438
Db 2465 CGACA-----GTTCCGCTTGTGGCGACCGGCTCTACCTGAGCTATT 2506
OY 1439 aacgagatcctcctcagatgtagaagcagcagcagcagcagcagcagcagcagcagc 1498
Db 2507 AATGGTGTGCTTGTGCTATGTTCAAAACCGCTGCTTGCAAACCTACCGTGGTACT 2566
OY 1499 gttgagatcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1539
Db 2567 GTCGAGCGTGGAGCTCTCAACCGCGGTAAACGTTGAGCGGTATGTTCTTCGGAAT 2626
OY 1540 -----tgacccctgttc 1550
Db 2627 ATGACTAGCCTTGCTAAGATTGAGAGAAATTTCCACTAACACTTTCAGACCTATTTC 2686
OY 1551 aatcaccctgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1601
Db 2687 ACATCATCTTGTGAGCTTCAAGGTATTTCTGATCTCCGCAACAGCGCGCACAG 2746
OY 1602 tcaagccttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1661
Db 2747 TCATGCTATACAGAG---TCCGGCTTCAAAAGAGTTGCTGCTGCTGCTGCTGCTG 2803
OY 1662 tctatgtagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1721
Db 2804 TGGTGTGTGAGGCTCTATTAAGCGGTATGATGATGATGATGATGATGATGATGATG 2863
OY 1722 ttcacatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1781
Db 2864 CTAATATCGGATAGCCTTTCC--TGGTATATACATGTTCCATTTGCCAATTTGATTCA 2921
OY 1782 cgaagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1841
Db 2922 CGAGGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2981
OY 1842 ctacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1901
Db 2982 TGGCACTGTTTCTGTTGACCTTATGAGAGAGCTTTGGCAGGCTCGCTTATGAACTCG 3041
OY 1902 agagtcgagaaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1961
Db 3042 CGAGTTCAGGCTCAGAGTGGCGACAGTTACAGGCTTACAGGCTGTTACGACGATTCAGAC 3101
OY 1962 gatgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1997
Db 3102 TATGCTGATATACAGACTTACGCGCACAGCTGACGA 3137

RESULT 11
AX195966
LOCUS AX195966 1422 bp DNA linear PAT 28-AUG-2001
DEFINITION Sequence 38 from Patent WO0151639.
ACCESSION AX195966
VERSION AX195966.1 GI:15386198
KEYWORDS
SOURCE
ORGANISM Micromonospora carbonacea.
Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae;
Actinomycetales; Micromonosporineae; Micromonosporaceae;
Micromonospora.
REFERENCE 1 (bases 1 to 1422)
AUTHORS Hosted,T.J., Horan,A.C. and Wang,T.X.
TITLE Evernimicin biosynthetic genes
JOURNAL Patent: WO 0151639-A 38 19-JUL-2001;
Schering Corporation (US)
FEATURES
SOURCE location/Qualifiers
1..1422
/organism="Micromonospora carbonacea"
/db_xref="taxon:47853"
1..>1422
CDS


```
/note="evic"
/codon_start=1
/transl_table=1
/protein_id="CAC6009.1"
/db_xref="GI:15386199"
/translation="MSRSLRDAQAQAAPASPNPHAGHAPVPSRVSTTVAVTP
TEPVPFRLTPVSRDIDYEIPRAVOVLIPLGLTPVATYAGSVGPTIARTG
RPRVITTYNGIDTANVHLHGHWATSDGHPMDLIPGSGKVDYDPMLOGATLWTH
DHTHAYEADHYRGLHGFYLLIDDEPAHHLRLIPAGKYDVPIMLRNAGFDSGLVGHF
DDRVTIILANGAOPYFEVAPRRYRFRLLNALKHVFRLNLGSEPLRTATGGLLPAP
TSHNELALSSGEREIVYIDFAHAGGSPVLYDGNPLTRFDVSSRAATDSRPVTL
RALPMSGTPTVETRVSMSEFMSARPPIALMGKFPDPLRVQVKGSTETEMVYNAD
TDPEFPHRLHLVTRFVLGRDGPAPEDAGLADIVYSPKSKIQVTFATPYLGS
QYVYHCHYLEHSSISGMAQLEVP"

BASE COUNT      219 a      575 c      431 g      197 t
ORIGIN

Query Match      4.6%; Score 95.8; DB 6; Length 1422;
Best Local Similarity 54.8%; Pred. No. 1.3e-12;
Matches 212; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

Oy 472 gcaatgaagatgtgtccgttcgtgaacagagagagagacacccctcccaacagcttc 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 299 GCACGGGCCGCCCGTGGGATCACTACACCAAGCGGCTCGACACCCGCAACGTGC 358

Oy 532 acttcaagagctcttctcgtcagctcccttgatgtgtgagctgaagacactaccagc 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 359 ACCTGACAGCGCGGACGACGCGCGGACACGAGGACGACGACGACGACGACGACG 418

Oy 552 ctggcgaatcaagaattactactaccaccaagaagcgtgcgcgactgtctgtgtacc 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 419 CGGCGGCTCGAGAGCTACGACCTACCGAACCTTCAGGCGGCGGACGACGCTGTG 478

Oy 652 atgaccatgcacatgcacacagcagagacactatctgtgtgtgtgtgtgtgtgt 711
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 479 ACGACCAACACCCAGCTTACGAGCGGACGACGCTTACGCGGAGCTGACGCGCTTATC 538

Oy 712 tgatcaagaacccgcgtgaagatgcctcgaacccccaagcagcgtgaagcttgata 771
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 539 TGATCGAGAGACCGCGCGGATCACTACGCGCTCGCGCGGCAAGTACGACGCGCA 598

Oy 772 tcccttggtctgtactgacacgaagatatacagcagaagcgaactctctctccaccatg 831
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 599 T---CATGCTGGCAACGCCCGCATGTCAGACTCCGCGCCCTCTCTGCGGCAACCCG 655

Oy 832 gagaggttccagctctcgtggtgagc 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 656 ACGACCGGCTCACATCTGCGCAACG 682

RESULT 12
AX195929      109519 bp      DNA      linear      PAT 28-AUG-2001
LOCUS      AX195929
DEFINITION      Sequence 1 from Patent W00151639.
ACCESSION      AX195929
VERSION      AX195929.1 GI:15386161
KEYWORDS
SOURCE      Micromonospora carbonacea.
ORGANISM      Micromonospora carbonacea.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Micromonosporinae; Micromonosporaceae;
Micromonospora.
REFERENCE      1 (bases 1 to 109519)
AUTHORS      Hosted,T.J., Horan,A.C. and Wang,T.X.
TITLE      Evernimycin biosynthetic genes
JOURNAL      Patent: WO 0151639-A 1 19-JUL-2001;
Schering Corporation (US)
FEATURES
source      1..109519
              /organism="Micromonospora carbonacea"
              /db_xref="taxon:47853"
BASE COUNT      15101 a      39910 c      39406 g      15102 t
```

```
ORIGIN

Query Match      4.6%; Score 95.8; DB 6; Length 109519;
Best Local Similarity 54.8%; Pred. No. 1.6e-12;
Matches 212; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

Oy 472 gcaatgaagatgtgtccgttcgtgaacagagagagagacacccctcccaacagcttc 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23046 GCACGGGCCGCCCGTGGGATCACTACACCAAGCGGCTCGACACCCGCAACGTGC 23105

Oy 532 acttcaagagctcttctcgtcagctcccttgatgtgtgagctgaagacactaccagc 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23106 ACCTGACAGCGCGGACGACGCGCGGACACGAGGAGGTACCGATGACGATCCGCG 23165

Oy 592 ctggcgaatcaagaattactactaccaccaagaagcgtgcgcgactgtctgtgtacc 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23166 CGGCGGCTCGAAGCTTACGACCTTACCGAACCTTCAGGCGGCGGACGACGCTGTG 23225

Oy 652 atgaccatgcacatgcacacagcagagacactatctgtgtgtgtgtgtgtgtgtgt 711
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23226 ACGACCAACACCCAGCTTACGAGCGGCGGACGACGCTTACCGGAGCTGACGCGCTTATC 23285

Oy 712 tgatcaagaacccgcgtgaagatgcctcgaacccccaagcagcgtgaagcttgata 771
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23286 TGATCGAGAGACCGCGCGGATCACTACGCGCTCGCGCGGCAAGTACGACGCGCA 23345

Oy 772 tcccttggtctgtactgacacgaagatatacagcagaagcgaactctctctccaccatg 831
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23346 T---CATGCTGGCAACGCCCGCATGTCAGACTCCGCGCCCTCTCTGCGGCAACCCG 23402

Oy 832 gagaggttccagctctcgtggtgagc 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23403 ACGACCGGCTCACATCTGCGCAACG 23429

RESULT 13
AX204969      3446 bp      DNA      linear      PAT 30-AUG-2001
LOCUS      AX204969
DEFINITION      Sequence 4 from Patent W0015180.
ACCESSION      AX204969
VERSION      AX204969.1 GI:15394251
KEYWORDS
SOURCE      Micromonospora carbonacea.
ORGANISM      Micromonospora carbonacea.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Micromonosporinae; Micromonosporaceae;
Micromonospora.
REFERENCE      1 (bases 1 to 3446)
AUTHORS      Stafla,A., Zazopoulos,E., Mercure,S. and Nowacki,P.
TITLE      Genetic locus for evernimycin biosynthesis
JOURNAL      Patent: WO 015180-A 4 02-AUG-2001;
Ecopia Biosciences Inc. (CA) ; Farnet, Chris (CA)
FEATURES
source      1..3446
              /organism="Micromonospora carbonacea"
              /db_xref="taxon:47853"
              misc_feature      3..1037
              misc_feature      /note="ORF 2 (positive strandedness) incomplete;
              C-terminus only (N-terminus undetermined)."
              1077..2231
              misc_feature      /note="ORF 3 (positive strandedness)"
              2242..3444
              /note="ORF 4 (negative strandedness) incomplete;
              C-terminus only (N-terminus undetermined)."
BASE COUNT      427 a      1148 c      1315 g      556 t
ORIGIN
```


BASE COUNT	1370 a	3694 c	3132 g	1335 t
CDS	5386..6657 /gene="aur2A" /note="Aur2A" /codon_start=1 /transl_table=1 /product="B-ketoacyl-ACP synthase-like protein" /protein_id="AAK61718.1" /db_xref="GI:14486277" /translation="MSRRVYVTIGVAVAPGGVTKREKPSLLTACKTARTALSLFDASPCV FRSLTAPADPNPLDGHGTAEEMRLRPAALFAYLSAEVNRDGIPEALDPPRQV TLGSAVCGTGLDVEYNVAVSRGGSTWQDHTLAPHLDFIVPESMAAEVARNVAG PVSWSVCTGSLDISGHAVELIREGSAADVITATEPLISPITVACDAIKATSPR DDPCTASRPEDATRNNGVFLVGGSAVLVLEDESARKRAHAYAIISGASCNVYHMT GLRADGVEMAEITVALDEARLDPDSADVITNAHSGTQRONDREHTAEFRSLGORAYE VPVSIKSMIGHSGLGAISSKVAASALAIENVNTPVNLHTPDLCDLDYTPLTAAE ORTAVLSVSGSGFGSGFAMILTRPRIGDAA"	6654..7916 /gene="aur2B" 6654..7916 /gene="aur2B" /note="Aur2B" /codon_start=1 /transl_table=1 /product="chain length factor-like protein" /protein_id="AAK61719.1" /db_xref="GI:14486278" /translation="MTTETARPGGTPLSPSTAAPVYTGIGVTAENGIGCTEAMAAVLRE GESIRVYGRFDATQYPAKLAGEVPEGPEAABVPSRLPQTDHMTRLALTNADBALD AGVNPALPDFSAGVITRSSAGVEFGQKELQALMSGQVYSAVSQSPWFYAVNTGQ ISINGNLPSPGVLVTEQAGLDVAQAARQLRKLGLVYGVGDSISCPWMAAHILA GGEISLSEDPARAFRRPESAEADGVAVAGGALLVLEDAARREGAEIYGTAVAGAST FDPAPSGSGAPRLOAAELARDAGLNPADVDVYFADAAGRDADRAEARLTALFGES YGVVTPATKMTGTGSLAGASLDLAAALLSLRDVYIPTVYVTGTGAEDCCPDLYSEVR RPELRRLAVLYARGRGFNMAVYRAAH"	7997..8500 /gene="aur2C" 7997..8500 /gene="aur2C" /note="Aur2C" /note="Aur2C: similar to acti dimerase from Streptomyces coelicolor" /codon_start=1 /transl_table=1 /product="oxxygenase-like protein" /protein_id="AAK61720.1" /db_xref="GI:14486279" /translation="MPPIPADPFARTMAHVPYVITVYVDRAQGRMGFTASSFSLSL LDPLVTEGPAKTASCHDAEYTGDRFLVNNVLAABHADIASSFASGDRKFTGSPMEPC EFGPLGDLDPATARIACTLHDVLDGGDHSILVGRVESVSAGAEPLVYNNRLFTPRGAS PALASTS"	8524..9441 /gene="aur2D" 8524..9441 /gene="aur2D" /note="Aur2D" /codon_start=1 /transl_table=1 /product="cyclase-like protein" /protein_id="AAK61721.1" /db_xref="GI:14486280" /translation="MTERSPVSAOLQOEVAEDVYAITOPDGSCMLNAGLITAGRPAL IDTATTEARAKALMDAVARYPHAPGFVINTHPGDHTEFGNRLFADEAVIITAHATP EMDLAAGTEALDMPDVGWGVLEPLTEFPEPTLHLGDIRAEVMIHLTAHTNDS VMEERRIYFAGDILVMSGARPFCLIGSSGLAVIEKRLALGPRTIYSGHGPGGTETL LDSEETATRYODLAGAGLACPLPEAKRAADLGPRAALLDSERIVPNLIRATVAAEAA GAEGHLIDISELSEMYDFHGGILPSCAA"

Query Match	3.9%;	Score 81.4;	DB 1;	Length 9531;
Best Local Similarity	54.1%;	Pred. No. 4.8e-09;		
Matches 216;	Conservative	0;	Mismatches 171;	Indels 12;
				Gaps 2

Oy 429 tacagctgcatgtccccaggaacctaccatcatcgtctccgtyggagcaatgaagtgttcc 488
 Db 1526 tTACGAGGCGACCTTCCCGCGCCGACCATCCGGGGACCCAGGGCCGTGAGTCTGTCGTC 1467
 Oy 489 cgccttcgtgaacagcggagagaaacactctctcccaacagcgtcactctgcagcgcctttcc 548
 Db 1466 CG-----GCAATATCAAGAGGCTCCAGGTGAGAACCCGGCGTCCACTGTGACGGGGGGCAC 1413
 Oy 549 tctcagctcccttgcattgattggtctggcctgaagagcaactaccagccttgcgagttacaagaat 608
 Db 1412 GTCCTCTCCGAGACAGACAGCGGCTCGCGATGAGACACCATGTCGCGGGCGGCGAAGTACG 1353
 Oy 609 tactactaccccaagaaggagcgtctcccgcatgctttgttaccatgatcatcattctcc 668
 Db 1355 TACCGCTTCCGACCAACACAGCGCGGGCGGTGCGTGTGTGTACCAAGAACCGCACACAC 1293
 Oy 669 atcacccgcgcgaagaacgcctacatctggctcaagctctgtgtctacatbatctcagaaccgcgt 728
 Db 1292 CTGGAGAGCGCCGGAAGAGTCTTCATGTGGCGCTGCACAGGGCTGTACTCTCTCACGACCAAC 1233
 Oy 729 gaggaatgccctgaacctctcccaagcgcgtacaggcgaagtttgatataccctcttggtctgact 788
 Db 1233 GAGGCCACAGCTCCGTTGTGCGAGCGGTGCC-----TAGAGAGTCCCGCTGTCATTCGGG 1179
 Oy 789 gcccaagcgtatacaagcagcagcgcctctctctccac 827
 Db 1178 GACGGCGGGTGTGAGAGCGGACGCAACCCGTGCTTACACC 1140

AP003233	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AP003233	137879 bp	DNA	linear	PLN 25-MAY-2001		
AP003233		Oryza sativa genomic DNA, chromosome 1, PAC clone: P0037C04.				
AP003233.3	GI:14090356					
		Oryza sativa (cultivar:Nipponbare) DNA, clone:P0037C04.				

ORGANISM *Oryza sativa*
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
Enharptoidae; Oryzaceae; Oryza.
1 (sites)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nippondare(GA3) genomic DNA, chromosome 1, PAC
clone: p0037C04
Published Only in Database (2001) In press
2 (bases 1 to 137879)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-296-38-7441, Fax:81-296-38-7468)
On May 15, 2001 this sequence version replaced gi:13365490.

Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI Nonredundant Protein database, nr (<ftp://ncbi.nlm.nih.gov/blast/db/>) and the cDNA sequence database at RCP. Protein homologues of the coding regions were searched against NCBI Nonredundant Protein database with BLASTX2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DBD accession no. and RCP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from Sp6 to T7 of the PAC clone. This sequence of P0037C04 clone has an overlap with P0445D12 (DBD):

AP003046) clone at the position 1 to 45,349 of 5' end. The sequence of this clone starts at the position 95,604 of P0445D12. This sequence of P0037C04 clone has an overlap with P0044F08 (DDBJ: AP002909) clone at the position 48,431 to 137,879 of 3' end. The sequence of this clone ends at the position 89,449 of P0044F08. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rtp.dna.affrc.go.jp/genomeseg.html>.

FEATURES

source	Location/Qualifiers
gene	1..137879 /organism="Oryza sativa" /cultivar="Nipponbare" /db_xref="taxon:4530" /chromosome="1" /clone="P0037C04" join(905..1139,1595..1860) /gene="P0037C04.1" join(905..1139,1595..1860) /note="hypothetical protein"
gene	/codon_start=1 /protein_id="BAB55515.1" /db_xref="GI:14090357" /translation="MAIWAARHGTAOAPRCGLSGWHGPTDAPCAVPAAGHRTTPMHGPIVGPGRAGPRAWAIVLPHKSLVLPQLSSSSKCYKHPRLULSKVYDERKROIHYCCRGRRPLSSORSQQPAVAVIMRPPSPMRIRIGLSPTPPQONDHIAAEVAVASA"
gene	complement(4956..8529) /gene="P0037C04.2" complement(4956..8529) /note="probably inactive due to frameshift(s) in CDS probably inactive due to stop codon(s) in CDS pseudogene, similar to gag-pol polyprotein"
gene	/pseudogene join(10530..10652,12160..12528) /gene="P0037C04.3" join(10530..10652,12160..12528) /gene="P0037C04.3" /note="hypothetical protein"
gene	/codon_start=1 /protein_id="BAB55516.1" /db_xref="GI:14090358" /translation="MODIPHEERLLDYREKSRDGDOKLPWGVQSMATDKTKPNLISITVVLVHLREDSRVLLINGGDVPGKRSLSVCAAGAEALTYMTAVSGAPGALSLASAGSVPRVRRWRYPTGGLFVPDAWMRASGGSVTVVHKLPPELEDTTAA"
gene	complement(15496..16914) /gene="P0037C04.4" complement(15496..16914) /gene="P0037C04.4" /note="contains EST C28475(C61239) similar to Oryza sativa chromosome 1, P0698A04.16 unknown protein"
gene	/codon_start=1 /protein_id="BAB55517.1" /db_xref="GI:14090359" /translation="MADPRPPCVLLERVVREVAAGLITSGCASRDADVAATENGWSMSRVOMGSAEMERRRAPSVPVAFADPQASSLNLHLPAPRTLLIGELISGTHKGIIVIAHRCVLLYDASNNHTAIPVPPSVFVPLGSALVSAAGDDDDYLIADIVTSCSRGILNPLEKATIFAVKNGEIOSIPLPLPPLCGPTFFHIDTAFSEAGTIFWVLLKGLICDELISPOGRILVFRLRICIDAHDRHCFSEFVHRSIGRVSGAIFKFLIGICEASCPANVKTWLSIPDFKMKKEETLITVGDIMVSSRONGLPHVPEPLGVNEDGIMVNLNVKEEPIRLNEFGDSLGIQVPRKANTIRRDMLONKVLSTKIKSKPTFRNFTMTFLASDFSAYIQRQVNLNLTATLATHTFLSDHPNKLVCFTLFTQTLFLIVTS"
gene	complement(19962..19079,19226..19636,19997..20204,20474..20575,20871..20948,21217..21321,21330..21552) /gene="P0037C04.5" complement(join(18962..19079,19226..19636,19997..20204,20474..20575,20871..20948,21217..21321,21330..21552)) /gene="P0037C04.5" /note="hypothetical protein"
CDS	
gene	/codon_start=1 /protein_id="BAB55518.1" /db_xref="GI:14090360" /translation="WVAALVGEVDADNTNGMGRRRPPYLIRGGGCGGRRPPDVAGGVVEEKASRSRGYPDPMPGPPDPAVGGGCGYRRIRIRLPSGCGGCGGRRSAERTQSGDEKREKRRKREDREEREKRLVDEVISDEALWSTITNRSOMGRGVAVALKRTSODRDSVAOKODPPEASSDGRVDGAGRAPSKIPARTETIWPPTPHASDARNGARROMIEPVGAHLGRLLVPAARLLPPAASRRPPPPPPADALARIIVVVAAPSSSSRLPSSSPSPAELAEFLSAGRARLISRRPPSSPSSPAPAAELSEFLSAGHARRRRPPSPSSPAPAVLSAASLASSCNOTLSKKRRCPRIYPTNQTHERVRRMSVAAAEGRR"
gene	join(25223..25413,25596..26220) /gene="P0037C04.6" join(25223..25413,25596..26220) /gene="P0037C04.6" /note="hypothetical protein"
gene	/codon_start=1 /protein_id="BAB55519.1" /db_xref="GI:14090361" /translation="MRPLRGLMKKNGKRTTMDHGTAKRTRSVAKKTEAGLCTDOODVEERTKMTYSIDSDIEACSPCNEPIGDIRCPLEKVLVLAASAPCKFRASGCMETVGFERTLSHASCITRAPCGCPDCTYIGLILYNIILDEHATDAVVAAGSLGTIVYHKSPPNVLVLRGSTRGNRVFLNLNGDVLSGRLSVSGVPPPTANCLKTELAADGPGCTGVILKLSAGTVPCVRLGEFNKAFLEVPDPSGSSDFTVSVTVLI"
gene	complement(join(27855..28310,29207..29638,29878..29901)) /gene="P0037C04.7" complement(join(27855..28310,29207..29638,29878..29901)) /gene="P0037C04.7" /note="hypothetical protein"
gene	/codon_start=1 /protein_id="BAB55520.1" /db_xref="GI:14090362" /translation="MMWLELRQANRRGRRAAEQVASEGPDGRRARPCGPPVARS SWRRPVSKTLASAERQAVYAASSPAMCOTEGRGGGGLQPLRRHAAKVASRLARCBEGARRRQQAQGHSGSGSNSTQSPVTVYEDLKLEIGLARARRRAAEQVAVASEGPDGRRARPCGPPVARSRRMPVSCKLSAERGQVAAVASEPMQGTGREGGCGPALQRLRRAAGKVASRLAREGARRRQAGRAAQQRRSGSGSNSTTOSOPTVYVIEDLKLEIGLIVSNTR"
gene	complement(join(30803..31498,31608..31763)) /gene="P0037C04.8" complement(join(30803..31498,31608..31763)) /gene="P0037C04.8" /note="hypothetical protein"
gene	/codon_start=1 /protein_id="BAB55521.1" /db_xref="GI:14090363" /translation="MDNTGTPKRTKMSENKAAQDCKDTTKLTYSIERDALECPICFVPEEDRVYMKNGHAAGSCYAKAMTMCPICIPINICRPLEKVLAAASAPCRSTSAVRLIRASGCTEIVATERRMHSAACPHAPCPDFGNNVGHLLYSHIDHEATDAAVATGCLRGCTGVTLHKSRPFVILLHRGGSRVLLINGDVLVSGRLSLVCTSPPLPNCBELKXIKELGAVSRAPGELGLSMGCTVPCVRLGDFDAKATLEVPDSTWSSGTISVTVHL"
gene	join(32915..33391,33775..34458) /gene="P0037C04.9" join(32915..33391,33775..34458) /gene="P0037C04.9" /note="hypothetical protein"
gene	/codon_start=1 /protein_id="BAB55522.1" /db_xref="GI:14090364" /translation="MSGSGVGRDRIAPAGTARRRRRDPPTGSPVYLEGSRWLSLPAVOPRRRRHDEGYSRRGRSRNSPPRRRRRSCRSRSRTSYSCDSRCRCHTBSPRDREDDEDDDDGSGYHRRPADNETVRLDIGADGDIFRCGCTAMLSPPITECANGDIVTCEKSTDDGGAVRCRCGTWELNARSAIGHILRCIRFACKNNRYGCPSLPRQMDMDEHLSCHDEPCFPIRRGFGAGADSLARHLTAHAGWGRLLVAAAGEAVPVQSPFLIRADGRIFLHLCSTRERGGGCTAASMCIRPDHYAGAEETVAVRACORLOM OAAVEGTSLRYGKMDAVOARVYVDDMLLRQDAVOVYRATSAAGANNNN"
gene	complement(join(35545..35965,38859..39226)) /gene="P0037C04.10" complement(join(35545..35965,38859..39226)) /gene="P0037C04.10" /product="putative Bowman Birk trypsin inhibitor"
CDS	

```

/protein_id="BAB55523.1"
/db_xref="gi:14090365"
/translation="MKVSTMAASMLFLVLVAGIVATTTTDSNIRLPSNGAGE
TDGNKOAKSRPMEDCDNTEMGYLKIYPRWRNDEVKOCAAACENCILQVPGAGEDV
VFVCDWMPTNPGPVCTPRPRMKCDNIVRLPERINPPFWOCDELEPGCFRCEA
CRDPGRPFPGHPLICDVFWDGDDPTSCAPSSSEMPWGPCCDIAVCTKSLPICHS
EVESCAACGOCSEWVDSWMLFVCRDSFTGDPGPRCTPEMHN"
complement(42298..43077)
/gene="P0037C04.11"
complement(42298..43077)
/gene="P0037C04.11"
/note="contains ESTs"

```

gene
CDS

```

Query Match      3.8%; Score 79.8; DB 8; Length 137879;
Best Local Similarity 61.1%; Pred. No. 1.3e-08;
Matches 129; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 609 taactactaccacaagcagcagctgcgcgcatgttggtaacatgacatgcatgtcc 668
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131795 TACACTTACCCCAACGCGAGTCCCGGCGTCTGTGTACCACGACCGCCCTCGGC 131854

QY 669 atcacgcgcgaagaagcctacatgtgctggtgtctacatgatacagagcccgct 728
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131855 CTACCCCGCGCACTCTCGCGGACTCTCGGCGCCTACGCTATCCGCAACCCGCGC 131914

QY 729 gagatgcccctgaacctccacagcggtacagcgagttgatatcccttggttctgact 788
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131915 GTGGAGGCGCGCTCGGCTCCCTCGCGGCGATGAGTTGACCGCGCTCCTCATGCTCGCC 131974

QY 789 gccaaagcatatacaagcagagcgactctct 819
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131975 GACCGAGCTTCTACGCGCGAGCTCCATCT 132005

```

Search completed: October 5, 2002, 20:28:59
Job time: 9074 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 20:28:34 ; Search time 448.02 Seconds
(without alignments)
8028.516 Million cell updates/sec

Title: US-09-656-640A-3
Perfect score: 2095
Sequence: 1 cagctcgctactactctc.....gaataagttgtgtcctaa 2095

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2095	100.0	2095	22	AAF82586
2	2095	100.0	2095	22	AAF30029
3	1946.2	92.8	1958	22	AAF30028
4	1944.6	92.8	1958	22	AAF82585
5	528.6	25.2	1959	14	AAO47790
6	517.4	24.7	2110	21	AAZ61243
7	460.4	22.0	1791	20	AAZ27601
8	460.4	22.0	1791	21	AAAS0019
9	460.4	22.0	1791	21	AAAS1314

10	458.8	21.9	1791	20	AAZ25727	Stachybotrys chart
11	300.4	14.3	2067	20	AAZ25735	Stachybotrys chart
12	300.4	14.3	2067	20	AAZ27609	Stachybotrys pheno
13	300.4	14.3	3676	20	AAZ25728	Stachybotrys chart
14	300.4	14.3	3677	20	AAZ27602	Stachybotrys pheno
15	300.4	14.3	3677	21	AAAS0018	Stachybotrys chart
16	300.4	14.3	3677	21	AAAS1313	Stachybotrys chart
17	288	13.7	2063	21	AAAS1316	Curtularia pallasc
18	288	13.7	2063	21	AAAS1316	Curtularia pallasc
19	283	13.5	2905	21	AAAS0020	Bipolaris spicifer
20	283	13.5	2905	21	AAAS1315	Bipolaris spicifer
21	95.8	4.6	109519	22	AAAS08693	Micromonospora DNA
22	87.4	4.2	858	21	AAAS0022	Americosporium atrum
23	87.4	4.2	858	21	AAAS1317	C glutamicum atrum
24	49.6	2.4	1533	22	AAH66018	C glutamicum codin
25	49.6	2.4	1614	22	AAH71220	Cornebacterium gl
26	49.6	2.4	349980	22	AAH68527	C glutamicum codin
27	47.2	2.3	523	21	AAA43363	Rat secreted expro
28	47.2	2.3	523	21	AAA43411	Gene encoding a su
29	44.8	2.1	10732	21	AAA10594	Propionibacterium
30	44	2.1	61313	23	AAAS5945	Human immune/haema
31	43.6	2.1	68356	22	AAK67283	Human immune/haema
32	43.6	2.1	68356	22	AAK83212	Human immune/haema
33	41.2	2.0	2012	22	AAH57385	Human skeletal mus
34	40.8	1.9	600	22	ABA59948	Human foetal liver
35	40.8	1.9	600	22	ABA28371	Probe #6837 for ge
36	40.8	1.9	600	22	AAK08219	Human brain expres
37	40.8	1.9	600	22	AAK34097	Human bone marrow
38	40.8	1.9	600	22	AAI39820	Probe #8506 used t
39	40.4	1.9	15772	22	AAK83220	Human immune/haema
40	39.4	1.9	2071	15	AAO58121	pH 2.5 acid phosph
41	39.4	1.9	2071	15	AAO56945	A. niger pH 2.5 ac
42	38.8	1.9	523	21	AAAB8173	P. meningitidis pa
43	38.8	1.9	20757	20	AAK20599	Polynucleotide seq
44	38.8	1.9	23457	22	AAK74866	Human immune/haema
45	38.8	1.9	31405	22	AAK67293	Human immune/haema

ALIGNMENTS

RESULT 1	AAF82586	standard; DNA; 2095 BP.
ID	AAF82586	
AC	AAF82586	
XX		
DT	18-JUN-2001	(first entry)
DE		Stachybotrys chartarum phenol oxidase B gene.
XX		
KW	Stachybotrys phenol oxidase B; spob; phenol oxidising enzyme;	
KW	detergent; paper production; pulp production; textile; food industry;	
KW	bleaching; ds.	
OS	Stachybotrys chartarum.	
XX		
XX		
FT	key	Location/Qualifiers
FT	CDS	77..2016
FT	FT	/tag= a
FT	FT	/product= "Stachybotrys phenol oxidase B"
FT	FT	77..267
FT	FT	/number= 1
FT	FT	268..319
FT	FT	/number= 1
FT	FT	320..869
FT	FT	/number= 2
FT	FT	870..952
FT	FT	/number= 2
FT	FT	953..1692
FT	FT	/number= 3
FT	FT	1693..1745
FT	FT	/number= 3
FT	FT	Intron

FT exon 1746..2013
PT /number= 4
PN WO200121748-A1.
PD 29-MAR-2001.
PE 06-SEP-2000; 2000WO-EP08840.
PX 22-SEP-1999; 99EP-0203120.
XX
XX (UNIL) UNILEVER NV.
PA (UNIL) UNILEVER PLC.
PA (HIND-) HINDUSTAN LEYER LTD.
PI Convents D, Doornink M, De Vries CH, Wang H;
XX
XX WPI: 2001-273462/28.
DR P-PSDB; AAB81505.
XX
XX New detergent compositions comprising a phenol oxidizing enzyme useful
PT in detergent or cleaning compositions, fiber treatment, processing,
PT finishing or production, paper and pulp production, or in starch
PT processing applications
PS
PS Example 3; Fig 3; 46pp; English.
XX
XX The present sequence encodes a Stachybotrys chartarum phenol oxidizing
CC enzyme. The invention relates to detergent compositions comprising one or
CC more surfactants and a phenol oxidizing enzyme having at least 688
CC identity to the Stachybotrys chartarum phenol oxidizing enzyme. Phenol
CC oxidizing enzymes may be used in the detergent, paper, pulp, textile and
CC food industries. They are used for preventing the transfer of dyes in
CC solution from one textile to another during detergent washing, or in
CC modifying the colour associated with dyes and coloured compounds having
CC different chemical structures, such as in pulp and paper bleaching,
CC bleaching the colour of stains on fabric and in detergent and textile
CC applications.
XX
XX Sequence 2095 BP; 437 A; 618 C; 510 G; 530 T; 0 other;

Query Match	Similarity	100.0%	Score	2095;	DB	22;	Length	2095;
Best Local	Similarity	100.0%	Precl.	No.	0;			
Matches	2095;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps
OY	1	cagctcgagctactactcgcgtctctctttgcaacaatcaacttaccatcgcttcctca	60					
Db	1	cagctcgagctactactcgcgtctctctttgcaacaatcaacttaccatcgcttcctca	60					
OY	61	attcaaacgatcatcaatgatcatgaccagaatctcgagagcgctgagctctgagcttgcgt	120					
Db	61	attcaaacgatcatcaatgatcatgaccagaatctcgagagcgctgagctctgagcttgcgt	120					
OY	121	gattcgcgagcagctctgctgcgatgcagatccgattgctgtgtatcatgaagagatgccttc	180					
Db	121	gattcgcgagcagctctgctgcgatgcagatccgattgctgtgtatcatgaagagatgccttc	180					
OY	181	cggtctcaccaaagagcagacgcagctgagtcctccctcccgagcttgaagaagtgacct	240					
Db	181	cggtctcaccaaagagcagacgcagctgagtcctccctcccgagcttgaagaagtgacct	240					
OY	241	ggcgcattcctctctctggaagcgcgcccaagtagtaagtaacttctatagcttaagaaagcca	300					
Db	241	ggcgcattcctctctctggaagcgcgcccaagtagtaagtaacttctatagcttaagaaagcca	300					
OY	301	agcttgctaactatgcgaataccgctcccaaccccaacacctcgagagagacattgtact	360					
Db	301	agcttgctaactatgcgaataccgctcccaaccccaacacctcgagagagacattgtact	360					
OY	361	acgagatgtagatctaggccctctctccacacagatcaccctgatacttgagccgagcaaca	420					
Db	361	acgagatgtagatctaggccctctctccacacagatcaccctgatacttgagccgagcaaca	420					

QY	421	tgtttgatacgaatgcaatgtccccaagacctaacaatcgtttccctgtgacatgaga	480
Db	421	tgtttgatacgaatgcaatgtccccaagacctaacaatcgtttccctgtgacatgaga	480
QY	481	gtgtgtcccgcttcgttgaacaagcgaggagaacaacctctccaaagcggtccaatgtgacg	540
Db	481	gtgtgtgtcccgcttcgttgaacaagcgaggagaacaacctctccaaagcggtccaatgtgacg	540
QY	541	gctcttctctcgaagctccctttgatgtgtgtgctgtagagacaataccagccgtgagagt	600
Db	541	gctcttctctcgaagctccctttgatgtgtgtgctgtagagacaataccagccgtgagagt	600
QY	601	acaagattactatactacccaacaaggcagagctgtcccgacatgtcttgtgtacatgacatg	660
Db	601	acaagattactatactacccaacaaggcagagctgtcccgacatgtcttgtgtacatgacatg	660
QY	661	ccatgtccatcaacgcgcgagagaagcctacatgtgttacaagctgtgtctacatgtaccag	720
Db	661	ccatgtccatcaacgcgcgagagaagcctacatgtgttacaagctgtgtctacatgtaccag	720
QY	721	accgggctgagagatgccccctgaacctccccaagcggtcagcgcgagtttgatatacccttgg	780
Db	721	accgggctgagagatgccccctgaacctccccaagcggtcagcgcgagtttgatatacccttgg	780
QY	781	tcttcaatgcacaagaacgataacaacgcagacgcaactctctcccaacaaatgagagagttt	840
Db	781	tcttcaatgcacaagaacgataacaacgcagacgcaactctctcccaacaaatgagagagttt	840
QY	841	ccaagctctcgtgggtgacgttatactcaagaatgtgtgaagttgagccattgagatgtctcaagtc	900
Db	841	ccaagctctcgtgggtgacgttatactcaagaatgtgtgaagttgagccattgagatgtctcaagtc	900
QY	901	ctaagaatcatcgaatgatacttgaatatctgtgcaatgcttaaccaagtgctacataccaagaagtgca	960
Db	901	ctaagaatcatcgaatgatacttgaatatctgtgcaatgctcttaaccaagtgctacataccaagaagtgca	960
QY	961	gacctgagcctatgtctcaacgltgcagccgcgcaaatgacacgctctccgctctcaacgctgtgc	1020
Db	961	gacctgagcctatgtctcaacgltgcagccgcgcaaatgacacgctctccgctctcaacgctgtgc	1020
QY	1021	cgctccaaagcctcttcgctctctgtatctctgtacatctcttgagagtttcaagagcaaacatcttc	1080
Db	1021	cgctccaaagcctcttcgctctctgtatctctgtacatctcttgagagtttcaagagcaaacatcttc	1080
QY	1081	cttccaaagttcatctgcgcgtgcagcggtgctgtctgtagggccctgtgtgacatgtacaactct	1140
Db	1081	cttccaaagttcatctgcgcgtgcagcggtgctgtctgtagggccctgtgtgacatgtacaactct	1140
QY	1141	gtatactctctatctgacgaagcgctgggagagttgtgtatcgaactcttccaaactctgcctgtagca	1200
Db	1141	gtatactctctatctgacgaagcgctgggagagttgtgtatcgaactcttccaaactctgcctgtagca	1200
QY	1201	gtccatcgatatactcgcaaacctctccgtgtgcagcggtctccggtgtcttgagccttgagtttga	1260
Db	1201	gtccatcgatatactcgcaaacctctccgtgtgcagcggtctccggtgtcttgagccttgagtttga	1260
QY	1261	taaacatgacaagaagltcatgtcgattcgttcgttgaagaagtcctttagtcgcgccgacaacttc	1320
Db	1261	taaacatgacaagaagltcatgtcgattcgttcgttgaagaagtcctttagtcgcgccgacaacttc	1320
QY	1321	tgaagttgcctgcgaacctctcgagatgttcccttcccgagggcggaacatgtgagcccgcg	1380
Db	1321	tgaagttgcctgcgaacctctcgagatgttcccttcccgagggcggaacatgtgagcccgcg	1380
QY	1381	aaaccccaatgatacgaagacttccacacttcgagccgtgtgataatgtgacagtgagacaatcaa	1440
Db	1381	aaaccccaatgatacgaagacttccacacttcgagccgtgtgataatgtgacagtgagacaatcaa	1440
QY	1441	cggagttacactctctcggagatgtcgaagaacggtcgtctcgcgcaatgtgtcccgcgacatgtc	1500
Db	1441	cggagttacactctctcggagatgtcgaagaacggtcgtctcgcgcaatgtgtcccgcgacatgtc	1500

```

OY 1501 tgaagatctggcagcttggagaacaactccaacggttggactcacccgttccatcaatc 1560
    |||||||
DB 1501 tgaagatctggcagcttggagaacaactccaacggttggactcacccgttccatcaatc 1560
OY 1561 cgttgacttcgcagatcccttctcgttccactcgcgttgagtcgagccttaagagctgc 1620
    |||||||
DB 1561 cgttgacttcgcagatcccttctcgttccactcgcgttgagtcgagccttaagagctgc 1620
OY 1621 tgggtcgaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1680
    |||||||
DB 1621 tgggtcgaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1680
OY 1681 cgcctccttccgcgaagtctcgccttcttaccactgttttccactcactcgttaacatct 1740
    |||||||
DB 1681 cgcctccttccgcgaagtctcgccttcttaccactgttttccactcactcgttaacatct 1740
OY 1741 acaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1800
    |||||||
DB 1741 acaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1800
OY 1801 gcttgcttccaatgtactcgttctggttactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1860
    |||||||
DB 1801 gcttgcttccaatgtactcgttctggttactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1860
OY 1861 cccatgtgagcctctctgtgagagcccccgcctcctcctcctcctcctcctcctcctcctc 1920
    |||||||
DB 1861 cccatgtgagcctctctgtgagagcccccgcctcctcctcctcctcctcctcctcctcctc 1920
OY 1921 ggtgacttcagcagcgttgcacatcactgcacgcgcatcagagagatggttagcttcaaccc 1980
    |||||||
DB 1921 ggtgacttcagcagcgttgcacatcactgcacgcgcatcagagagatggttagcttcaaccc 1980
OY 1981 taagccacagcgtgagatgagatccgcgtgagagtaataatgtgtgtgtgtgtgtgtgtgt 2040
    |||||||
DB 1981 taagccacagcgtgagatgagatccgcgtgagagtaataatgtgtgtgtgtgtgtgtgtgt 2040
OY 2041 tggacagcagatataagatattttagaataactgtgaataatgtgtgtgtgtgtgtgtgt 2095
    |||||||
DB 2041 tggacagcagatataagatattttagaataactgtgaataatgtgtgtgtgtgtgtgtgt 2095

RESULT 2
AAF30029
ID AAF30029 standard; DNA; 2095 BP.
AC
XX AAF30029;
DT 23-APR-2001 (first entry)
XX
XX Stachybotrys chartarum phenol oxidase B spob gene.
DE
XX
XX Phenol oxidizing enzyme; phenol oxidase B; spob gene; bleach.
KW pulp; paper; textile; detergent; ss.
XX
XX Stachybotrys chartarum.
OS
XX
XX Key Location/Qualifiers
FH 77..2016
FT CDS
FT /tag= a
FT /note= "contains introns"
FT exon 1..267
FT intron /tag= b
FT 268..319
FT intron /tag= c
FT exon 320..869
FT /tag= d
FT 870..952
FT intron /tag= e
FT exon 953..1692
FT /tag= f
FT intron 1693..1745
FT /tag= g
FT exon 1746..2095

```

```

FT /tag= h
XX
XX US6168936-B1.
PN
XX
XX 02-JAN-2001.
PD
XX
XX 22-SEP-1999; 99US-0401476.
PF
XX
XX 22-SEP-1999; 99US-0401476.
PR
XX
XX (GENEV ) GENENCOR INT INC.
PA
XX
XX Wang H;
PI
XX
XX WPI: 2001-136715/14.
DR
XX
XX P-PSDB; AAB20097.
DR
XX
XX New phenol oxidizing enzyme, also useful in the detergent, paper and
PT pulp, textile or food industries, especially in modifying the colour
PT associated with dyes and coloured compounds, as well as in anti-dye
PT transfer applications
PT
XX
XX Claim 5; Fig 3; 23pp; English.
PS
XX
XX The present sequence is that of the Stachybotrys chartarum MUC1.38898
CC spob gene encoding phenol oxidase B (see AAB20097). The gene was
CC isolated by PCR using primers (see AAF30030-31) based on 2 peptide
CC fragments of the isolated enzyme. The invention provides phenol
CC oxidizing enzymes such as phenol oxidase B, or enzymes with at
CC least 68% identity to it, nucleic acids encoding them, expression
CC vectors, filamentous fungus and yeast host cells, and methods for
CC the recombinant production of the phenol oxidizing enzymes. The
CC enzymes are useful for bleaching of pulp and paper, fabric stains,
CC and in detergent and textile applications. They show optimal
CC activity at pH 5-11, 7-10.5 or 8-10, and 20-60 or 20-40 degree C.
XX
XX Sequence 2095 BP; 437 A; 618 C; 510 G; 530 T; 0 other;

Query Match 100.0%; Score 2095; DB 22; Length 2095;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2095; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 cagctcgcgttactactcctcgcgttctccttgcgaatacaatacctacatcgttccctca 60
    |||||||
DB 1 cagctcgcgttactactcctcgcgttctccttgcgaatacaatacctacatcgttccctca 60
OY 61 attcaaacgatcaaatgatcatgacccaagctatcgagcgcgtgtgtgtgtgtgtgtgtgtgt 120
    |||||||
DB 61 attcaaacgatcaaatgatcatgacccaagctatcgagcgcgtgtgtgtgtgtgtgtgtgtgt 120
OY 121 gatcggcgagcactctgtgtgatgcagatcgttctgtgtgatcgtgacagatcgtcttc 180
    |||||||
DB 121 gatcggcgagcactctgtgtgatgcagatcgttctgtgtgatcgtgacagatcgtcttc 180
OY 181 cgtctcaccagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240
    |||||||
DB 181 cgtctcaccagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240
OY 241 gccgatccctctctgaaagcgcccaagtagtaattacattataggcttaagcaagca 300
    |||||||
DB 241 gccgatccctctctgaaagcgcccaagtagtaattacattataggcttaagcaagca 300
OY 301 acgttgttaattatgtcagatcagtcgtccccaacccaacccaacccaacccaacccaac 360
    |||||||
DB 301 acgttgttaattatgtcagatcagtcgtccccaacccaacccaacccaacccaacccaac 360
OY 361 acgagatgagataggccctctccaccagatctacccttatcttgagccgagca 420
    |||||||
DB 361 acgagatgagataggccctctccaccagatctacccttatcttgagccgagca 420
OY 421 tggttgatacagatgcatgtccccaagcactacatcatcgttccctgtgtgacatgaga 480
    |||||||

```

```

Db 421 tgggtggaatacga tggcaltgcccaggaactacatcagltccctgctggcactgaga 480
Qy 481 ggtgtgtccgcttcgtgaacagcgagaaacacactctccacacagcgctccacttgaca 540
Db 481 ggtgtgtccgcttcgtgaacagcgagaaacacactctccacacagcgctccacttgaca 540
Qy 541 gctcttctctcgaagctccctttgagltgtggctgtaggaacaactaacagcctggcgagt 600
Db 541 gctcttctctcgaagctccctttgagltgtggctgtaggaacaactaacagcctggcgagt 600
Qy 601 acaaggaattactactaccccccaacaggaagcgctgcgcgcagctgtgttgggtaccatg 660
Db 601 acaaggaattactactaccccccaacaggaagcgctgcgcgcagctgtgttgggtaccatg 660
Qy 661 ccatgtccatccacccgcgcgcgaagcgtctacatgtagagctgtgtgtctacatgacatc 720
Db 661 ccatgtccatccacccgcgcgcgaagcgtctacatgtagagctgtgtgtctacatgacatc 720
Qy 721 acccggtctgaggaatgacctgaaactcccccagcgctacggcgagtttgatcccccctgg 780
Db 721 acccggtctgaggaatgacctgaaactcccccagcgctacggcgagtttgatcccccctgg 780
Qy 781 ttcttactgcgaagcagataacacgcagacgacactctcttccacacaaatggagagttt 840
Db 781 ttcttactgcgaagcagataacacgcagacgacactctcttccacacaaatggagagttt 840
Qy 841 ccagcttctgggtgtagctgtatctcaagtgtgaagttgagcccatgtagatgtgttcacatc 900
Db 841 ccagcttctgggtgtagctgtatctcaagtgtgaagttgagcccatgtagatgtgttcacatc 900
Qy 901 ctagaagatcgaatgataaattgtgcatgctctaaacagctgctatacagaacgggtca 960
Db 901 ctagaagatcgaatgataaattgtgcatgctctaaacagctgctatacagaacgggtca 960
Qy 961 gcttggcctcattagctcaaaatgacggcgacgaatgacggcttcgcgcttccatcaacgctgc 1020
Db 961 gcttggcctcattagctcaaaatgacggcgacgaatgacggcttcgcgcttccatcaacgctgc 1020
Qy 1021 cgtctacagctctcttcgctctgtatctgtactacactctgagagatccagagacagacttc 1080
Db 1021 cgtctacagctctcttcgctctgtatctgtactacactctgagagatccagagacagacttc 1080
Qy 1081 ctctcagatcattgcgcgtctgacggtgtgtctgtctgagggcctgtgtgacactgacaactct 1140
Db 1081 ctctcagatcattgcgcgtctgacggtgtgtctgtctgagggcctgtgtgacactgacaactct 1140
Qy 1141 gtacatctctatggcgcgcgcgtctggaggtgtgtatcgagctctccacactctgcgcgcga 1200
Db 1141 gtacatctctatggcgcgcgcgtctggaggtgtgtatcgagctctccacactctgcgcgcga 1200
Qy 1201 gtccatcgatatacgcgaacactctcgtgtcgtacggtctcgtgtgtgagcccgagtttga 1260
Db 1201 gtccatcgatatacgcgaacactctcgtgtcgtacggtctcgtgtgtgagcccgagtttga 1260
Qy 1261 taacactgacaaggtcgaatgagatctgtctgtatgtaagctcttgagtcgcccagacattc 1320
Db 1261 taacactgacaaggtcgaatgagatctgtctgtatgtaagctcttgagtcgcccagacattc 1320
Qy 1321 tgaagtgctgcgcgaacactcccgagatgttcccttcccccgcgcgcgcgcgcgcgcgcgcgc 1380
Db 1321 tgaagtgctgcgcgaacactcccgagatgttcccttcccccgcgcgcgcgcgcgcgcgcgcgc 1380
Qy 1381 aaaccccaactgatacgaagacttccactctgcgcgcgtgctaaatggacagttgacaatcaa 1440
Db 1381 aaaccccaactgatacgaagacttccactctgcgcgcgtgctaaatggacagttgacaatcaa 1440
Qy 1441 cggagttaccctctcggatgtcgaagaacgcgtctgtctcgcgaatgtgcccgcgcgcgcgcgcgc 1500
Db 1441 cggagttaccctctcggatgtcgaagaacgcgtctgtctcgcgaatgtgcccgcgcgcgcgcgcgc 1500
Qy 1501 tgaagatcggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1560
Db 1501 tgaagatcggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1560

```

```

Qy 1561 cgttgaattccgaagttccttctcgtcttccacttgcgcgcgtgagagtcgaagccttatgaggtgc 1620
Db 1561 cgttgaattccgaagttccttctcgtcttccacttgcgcgcgtgagagtcgaagccttatgaggtgc 1620
Qy 1621 tgggtcgaagatgtgtctgctgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1680
Db 1621 tgggtcgaagatgtgtctgctgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1680
Qy 1681 cgtctcttccgcgtgaagttctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1740
Db 1681 cgtctcttccgcgtgaagttctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1740
Qy 1741 acaagttggtctacatgctgtgacgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1800
Db 1741 acaagttggtctacatgctgtgacgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1800
Qy 1801 gctgtcttcaatgtcaactgtctcgtgtgactatgtgctaacactaacacgcagatltgac 1860
Db 1801 gctgtcttcaatgtcaactgtctcgtgtgactatgtgctaacactaacacgcagatltgac 1860
Qy 1861 cccatggagcctctctgtagagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1920
Db 1861 cccatggagcctctctgtagagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1920
Qy 1921 ggtgacttccagcagctgttccatccactgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1980
Db 1921 ggtgacttccagcagctgttccatccactgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1980
Qy 1981 tacgcccagcgtcgtatgataatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2040
Db 1981 tacgcccagcgtcgtatgataatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2040
Qy 2041 tggacagcagatatagctattttagaataacttgaaatgttgggtgctaa 2095
Db 2041 tggacagcagatatagctattttagaataacttgaaatgttgggtgctaa 2095

RESULT 3
AAF30028
ID AAF30028 standard; DNA; 1958 BP.
XX
AC AAF30028;
XX
DT 23-APR-2001 (first entry)
XX
DE Stachybotrys chartarum phenol oxidase B spob gene.
XX
KW Phenol oxidizing enzyme; phenol oxidase B; spob gene; bleach;
KW pulp; paper; textile; detergent; ss.
OS Stachybotrys chartarum.
FH
FT Key
FT CDS
FT 13..1952
FT /*tag= a
FT /note= "contains introns"
FT 1..203
FT /*tag= b
FT 204..255
FT /*tag= c
FT 256..805
FT /*tag= d
FT 806..888
FT /*tag= e
FT 889..1628
FT /*tag= f
FT 1629..1681
FT /*tag= g
FT 1682..1958
FT /*tag= h
FX
PN US6168936-B1.

```


OY	603	aaggatctactacaccccaacgaagcgagcgctgcgcgcgaatgctttgtttacatgatcaatgcc	665
Db	551	aaagactatactactccccaataagacatctgtccgtacacccatagtacaacagatcaatgcct	590
OY	663	atgtccataccaacgcgcgaagaacgcctccatactagtgttcagagctgtgtctacatagtatccaagac	722
Db	591	atgcatactactgtcttgagaacgcctccacgcctacgcgcgcgcgtgtctctactatgtctcaatgac	650
OY	723	ccggctgtagagatgtgcccctggaacccctccacgagcgctaaagcgagtttgatatcccccttggt	782
Db	651	ccaaacgcgaagagcgcctcccaactctgcacaagctgagatagtgagatctgcgatactccaatgac	710
OY	783	ctgagcttccaagcgaatacaacgcgcgaagacgcacctctcttccacaacatgtgagaagttcttc	842
Db	711	ctcaagctccaagcaataatactacgcgaacagcgaactgtgtacacataatgtgagagctgtacac	770
OY	843	agctctcgggggtgacgttatctatcaagctgagatgagttgagcccatgtagatgctcagaacct	902
Db	771	tcattcttgagggtatgtatgaattctacgt-----	796
OY	903	agaaagtatcgatgtatgaataatgtgcattgtctcaacagatgtcatcacagaacggtacgc	962
Db	797	-----gaacggtccaac	807
OY	963	cttgagctatgctcacaagctgcagcgcgcgcgaagtaacgcgtcttcgcctccataacgcttcgcg	1022
Db	808	ccgtgaccttccaagaagacgtttgagacgcctccgcacaatatgcattccgcctctccgtatgcgcgacg	867
OY	1023	tctcaagcctcttcgcgcctgtatctctacaccccttcgaagatctcaagacccaacatctccct	1082
Db	868	ttctcgtcctcttcgcgccttactctgtcgtatcatctgatactatgcatactgcacactgcgttcgctt	927
OY	1083	tccaggtcatatgcgcgcctgcacgctgtgtctcgtcttgagagcgccctgttgacactgtacactctgt	1142
Db	928	tcaaggtatattgtccctccgacttctgcgtctctctcttgaaacacccctccgataacagctgtctgt	987
OY	1143	acatctctatagccgcgaagcgcgtcttgagaggttgatactgcactctccacactctcgtctgcgcgt	1202
Db	988	acattctccaatgcgcgaagcgttcaacgaagctgtgcttttgactctccgcactatgctctgcgcgaaga	1047
OY	1203	ccatctgatalccgcgaacactctccctgtgag-----ctgacgctctcgtgtgttgagccctgtgattg	1259
Db	1048	ctattgaaactccgcgaacactctggcgttgatgacatgctgcgcgcatactgcgaatacgaactatg	1107
OY	1260	ataacactgtgaaaggtatcatctgcatatgcgttgatgtgataagatctcttaagtctgcgcgcgaacct	1319
Db	1108	acaacacgcgaataaaggtcatatgcgtttcttggttgaaacagacacaacatccgcagaataact	1167
OY	1320	cttgaggtgtgcctgcgcgaacactccgcgaatgtctctctccgcgagcggcgaactgtgacccgcg	1379
Db	1168	cagttgttccctgcctcaaacctctgcgtatgttccctctccctctccacacaaacacacccccc	1227
OY	1380	caaaacccaactgatagcagagacttccacacttcgcgcgcgtgtctaaatgtgacagatgtgacatatca	1439
Db	1228	gaca-----gttccgcgtttgtgtccgacgcgctgcctactactgactatla	1269
OY	1440	acggagattactctctgcgtatgtctgagaacaacgttctgtctccgcgaatgtgcocccgcgaactg	1489
Db	1270	atggtctgtgcttcttgcgtatgtgttccaacaacgcgtctgcgtctgtcaaaagctaacccgtttgttatg	1399
OY	1500	ttgagatctgcgcgcgtgagaatacaactcccaacgcaggtttgagctccacccgtttcaactatcac	1559
Db	1330	tcgagcgtgtgagatctatcaataacgcgcgtgaacoggtttgagcgcgaactctatccaatcaatca	1389
OY	1560	tcgttgactccgcgaatctcttctcgttc-----cacgtcccggtgagatctgagacctc	1610
Db	1390	ttgtgcgaacttcaaggtatattctctgcacttccgcgcgaacaacagcgcgcgaagtcatgcatat	1449
OY	1611	atgagcgctgcctgcgttccaagaagtgtgtgtctgcgtctgcgtctgtctgtatgtgtgtatgtgt	1670
Db	1450	acgag-----tcgcggtctccaagaagatgtctcgtcgtgtgtctgcgcgcgttgaataactgtgtgtgt	1506

QY	1671	aggccacatgacgtctcttcgccgtaagtcttcgctttactactaactggtttactcat	1730		
Db	1507	aggccatcaagcgcccttccc-----	1528		
QY	1731	gttaacatctacaagtgtgtgtctacatgtttgacgtccacaacgtatccagagacca	1790		
Db	1529	-----tgtgtgtataatgttccattgtccatgtgcacaatattgtatccaagagatca	1573		
QY	1791	cgacatgatgctgtctcttccaatagtccaatgttctcggttgaactatggtctacaataccaga	1850		
Db	1574	cgatctgatgtgtgctctttaaagccacgcgtctctgcagatlatggtctataaagccactgt	1633		
QY	1851	gttcaattacccccatggagagccctctctgtgaagcccccgcctctccctcggagaagttcga	1910		
Db	1634	tttcgttgaaccttaaggagagagcttcttggaagcctctgcctctatgaactcggcgaagttcca	1693		
QY	1911	gaattgctcgggttgacttcaagcagagctgtgacctcactgaccgcatccagagatgcttag	1970		
Db	1694	ggctcagagctggccaggttccaggtctcagctgttacttgatgagctatcagactatgctcga	1753		
QY	1971	cttcaacccctaacgcccacgagctgatga	1997		
Db	1754	atacagaccttaccgctgcagctgacga	1780		
RESULT 6					
ID	AAZ61243	AAZ61243 standard; DNA; 2110 BP.			
AC	AAZ61243;				
XX					
DT	30-MAY-2000	(first entry)			
XX					
DE	DNA encoding a phenol oxidising enzyme.				
XX					
KM	Phenol oxidizing enzyme; fungus; redox reaction; detergent; bleaching;				
KM	fabric; pulp; paper; decolourisation; plant-derived food product;				
KM	coloured compound; porphyrin; tannin; polyphenol; carotenoid;				
KM	anthocyanin; Malliard reaction product; ss.				
XX					
OS	Acremonium murorum.				
XX					
FI	Key	Location/Qualifiers			
FT	CDS	135..1943			
FT		/tag= a			
FT		/product= "phenol oxidising enzyme"			
XX					
PN	WO200005349-A1.				
XX					
PD	03-FEB-2000.				
XX					
PF	13-JUL-1999;	99WO-EP04922.			
XX					
PR	21-JUL-1998;	98EP-0202454.			
XX					
PA	(UNIL) UNILEVER NV.				
PA	(UNIL) UNILEVER PLC.				
PA	(HIND-) HINDUSTAN LEVER LTD.				
XX					
PI	Convents D, Gouka RJ, Van Der Heiden M, Swarthoff T, Verrips CT;				
XX					
DR	WPI: 2000-195101/17.				
XX	P-PSDB: AAY69204.				
PT	Phenol-oxidizing enzyme from Acremonium, used in detergent compositions				
PT	for bleaching stains on fabrics				
XX					
PS	Example 2; Page 37-40; 45pp; English.				
XX					
CC	The present sequence encodes a phenol oxidizing enzyme from the				
CC	fungus Acremonium murorum. The enzyme has the CBS accession number				
CC	157.72. The enzyme catalyses redox reactions and is specific for				
CC	molecular oxygen as the electron acceptor. The phenol oxidising enzyme				

CC is specifically used in detergents for bleaching strains on fabrics,
 CC but also for bleaching pulp and paper and for decolourisation of
 CC plant-derived food products. The enzyme has a pH optimum in the
 CC alkaline to neutral range and can bleach a wide variety of coloured
 CC compounds, e.g. porphyrins, tannins, polyphenols, carotenoids,
 CC anthocyanins and Mallard reaction products.

XX Sequence 2110 BP; 400 A; 732 C; 556 G; 422 T; 0 other;

Query Match 24.7%; Score 517.4; DB 21; Length 2110;
 Best Local Similarity 61.4%; Pred. No. 1e-135;
 Matches 1041; Conservative 0; Mismatches 476; Indels 178; Gaps 6;

```

QY 321 accgtcccaaccccaactgagagagacattgttactacagagatgagatagccc 380
    ||||| ||||| ||| ||| ||| ||||| ||||| ||||| |||
Db 390 accgtcccaaccccaactgagagagatcgagactacatgagatcgagatcaagcac 449
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 381 ttctcccaacagatctacactgtactgtgagccgccaacatggttgatagcatg 440
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 450 ttctctgagagagctctccctgacctggcgctgactgactgtggtctagcagcgc 509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 441 tcccaagacatacactcgtctcctcgtgcaactgagagtggtgctgcgtgtaac 500
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 510 tcgcgcgtctactactcactccagagagagagcgctgctgcgtctcgtgtaac 569
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 501 agcgagagagacacccctcccaacagcttccactgtcagcgctcttctcagctccc 560
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 570 a-----aggcacagctcgagagctccatccatccatccatccatccatccatcc 623
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 561 ttgatgtgtgagctgagacactacacagctgagagatgagatgactactacc 620
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 624 tgggaagcgctgagctgagatgacacacacccggcgagacagactactactacc 683
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 621 aacagcagagctgcccagatcttgggtacatgacatgacatgacatgacatgac 680
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 684 aatgcgacagcgagcgctctcgtctgacacatgacatgacatgacatgacatgac 743
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 681 aagcgctacatgagtgagtggtgtctacatgacatgacatgacatgacatgac 740
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 744 aatgcctacacagcgagctgagtggtctacatgacatgacatgacatgacatgac 803
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 741 aacgtcccaacgagctgacagctgagatgacatgacatgacatgacatgacatg 800
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 804 ggtctgcacagcgatagtagagtagacatccctcactcactcactcactcactcact 863
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 801 aacgacagagcagctctctctccacacatgagagagttccacgctctcggtgacgt 860
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 864 aactccgacgtaactctctctcactcaagggtagagacagagctctctggtgagatg 923
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 861 attcaagtggttaagtgagcccatgtagatgactcactagaagtacgtatgata 920
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 924 atccaggt----- 931
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 921 aattgtgcatgtctaaccagtgatatacaagaacggtcagccttgctatgtcaacg 980
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 932 -----gacgagagcccggtgacctactctgacg 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 981 tgcagcgcgcgaagtaacgcttccgcttccctcaacgctgcgtctcaacgcttccgctc 1040
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 ttgagcccgagagtaactcttccgcttgcgtcgcagcgcgcgtctcctctcgcgc 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1041 tgtacttcttaactctgagagatcagagacagacttccctccagagtgatgacgctg 1100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 tctactttgctgacacgcgcgagatgagagacacccgcacatccggtccagtgatccgcctcg 1080
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1101 acggtgtgctgttgagccgctgttgacactgtacactgtacatctatgaccgagc 1160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 acctgtgtccctcgagagcgctcagacccctccacagctcgtctcctcatgtcgagct 1140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1161 gctgagagagtggtatcagacttccacactgtgctggtccagctccatccatccagcacc 1220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 gctacagatcactcctcgtacttctcgtactcgtcagagcgaagaccatccagctcgcgaacg 1200
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 1221 ttctgtgtgtagcgggtctcgtgtgttgagccctgattgtgataacactgacaagtgatc 1280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1201 agcccggtgtcgcgcgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1260
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1281 gatt---cgtcgttgatgaaagtcccttgagtcgcgcgcacacttctgagtgctgcgaac 1337
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1261 gcttcaacgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1320
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1338 tccgagatgttctcttcccgagggcggcacactggagcccgacaacccacatgagagc 1397
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1321 tccgtgacgtgctcttccc-----tccagcacctcgcacacacatcgacc 1365
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1398 agacttcaacttcgcgcgcgtgtaabgagactggagacacacacagagatctctcgcg 1457
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1366 actcgttccgcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1425
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1458 atgtcagaaacgctgtctcgcgaatgtgcccgcgacacatgtgtgagatctgagactg 1517
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1426 actcgtgaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1485
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1518 agacaactccaaaggttgagtaaccctgttcaacatccactcgttgaactccgagctc 1577
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1486 cgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1545
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1578 ttctc-----gttccactgcgcgcgttgagtgagtcgcgcgtttagaggtcgtg 1619
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1546 tgtctgcacaggtgtcgtgtaacggtgtgctacgacccgtgtgtgctgagccgtacagatcgtg 1605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1620 ctgctcgaagagatgtgtgtcgtgctgctcgtgtgaggtgtctatgtgtgagccacat 1679
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1606 ctgctcgaagagagctgtgtgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1665
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1680 acgctccttcccgtaagtctcgtccttctaactaactgtgttcaactatgtaacatc 1739
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1666 acgctccttgcgcgc----- 1679
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1740 tacaagtgtgtctatagtgtgactgacatgacacacacacacacacacacacacacacatgata 1799
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1680 -----ggtcgtgtacatgttccacatgacacacacacacacacacacacacacacatgata 1732
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1800 ggtcgttctcaatgtcactgtctcgtgtgactatggtctacacacacacacacacacacacacacacatgata 1859
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1733 ggtcgttctcaatgtcactgtctcgtgtgactatggtctacacacacacacacacacacacacacacatgata 1792
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1860 cccacgtgagcctctctgtgagggcccgccctcctcctcgtcgtgagaggttcgagaaatgctc 1919
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1793 cccacgtgagcagagatctccgcgcgaagccgtacgttgagacacagacgttcgaggtgcgcga 1852
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1920 ggtgtgactcagagagctgtgcacatgacgcgcacatcagagagatgctgactcaacc 1979
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1853 gaacgcttccagacacagacagatcgagcgaggtccaggtccagatgagtcgactcgtcc 1912
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1980 ctacgcacaggtcga 1994
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1913 ctacgacacacgcga 1927
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 7

AA227601
 ID AA227601 standard; DNA: 1791 BP.

AA227601:

16-DEC-1999 (first entry)

Stachybotrys phenol oxidase coding sequence.

phenol oxidase; enzyme; coloured compound; dye transfer prevention;

fabric washing; stain bleaching; anti-dye transfer; detergent; ss.

Stachybotrys chartarum.

PN W09949020-A2.
 XX 30-SEP-1999.
 XX 23-MAR-1999; 99MO-US06327.
 XX 24-MAR-1998; 98US-0046969.
 PR 22-DEC-1998; 98US-0218702.
 PR 22-MAR-1999; 99US-0273957.
 XX (GENE) GENENCO INC.
 PA Amoy A, Wang H, Dhase P, Lambrechts-Kongvaux A, Wang C;
 PI WPI; 1999-591088/50.
 DR P-PSDB; AAY3992.
 XX Novel enzyme for modifying coloured compounds used to prevent
 PT dye-transfer -
 PS Claim 21: Fig 5; 64pp; English.
 CC This sequence encodes the Stachybotrys chartarum phenol oxidase enzyme
 CC of the invention. The invention is used to modify a coloured compound and
 CC prevent dye transfer during fabric washing, or for stain bleaching or
 CC anti-dye transfer. It is useful in the detergent, paper and pulp, textile
 CC and food industries.
 SQ Sequence 1791 BP; 380 A; 551 C; 451 G; 409 T; 0 other;

Query Match 22.0%; Score 460.4; DB 20; Length 1791;
 Best Local Similarity 59.3%; Pred. No. 1.2e-119;
 Matches 996; Conservative 0; Mismatches 521; Indels 163; Gaps 6;

QY 321 accgtcccaaccccaactgtgagagagacattgttactacagatgtagtggccc 380
 DB 241 atcattaccaacccgttgcacggcaagagacattgttactatgagatcgaatcaagcca 300
 QY 381 ttctccacacagatctacactgtactgtgagccgcaacatgtttagatagctggcatg 440
 DB 301 ttccagcaagagtttaccacaccttgcgcctgcacactctcgtgctagctagtgccg 360
 QY 441 tcccagacactacatcactcgtctcctcgtgacactgagagtgctgcgctcgtgaac 500
 DB 361 agccctgtctacttacttcaatgttcccaaggagacagagactgtagttacatcaac 420
 QY 501 agggagagagaaacacctctcccaacagcttccacttgcacggctcttctctgagctcc 560
 DB 421 a-----atgcaacccgtgagagacatcgtctcacttgcacggctcccatcgctgccc 474
 QY 561 ttgtatgtctgtgagagacactacacacacgtgagctgagatagatattactactccc 620
 DB 475 ttcgatgttgcctgtaagatgtgaccttccctgcgcgagacagatatactacttccc 534
 QY 621 aacagcagagctgcgcgcactgtctgttaccatgacacatgcatctccatcaccgcgag 680
 DB 535 aactaccatcccgcccgcttctgttaccatgacacacgcttcttcaatgaagactgtgag 594
 QY 681 aacgcttaacatggttagagctgtgtctatcatgatccagaccggctgtagagcttcg 740
 DB 595 aatgcctacttctgtagcgtgcgcctacatatacaacagacagagctgagatgtcttc 624
 QY 741 aacctcccccagctagcagatgttgaatlatcccttggctttagctcgaagcgatag 800
 DB 655 ggtcttcctcctagtcagatgagctgcgatatacctctgtactctgagcgcaagtaactat 714
 QY 801 aacgcagcagcactctctctccacacatgagagagttccagactctcgtgggtgagct 860
 DB 715 aacgcagcagcactctctctccacacatgagagagagttcgtgggtgagagctgct 774
 QY 861 attcaagtgtagagtgagcccatgtgagatgtctcagatccctagaagtatcgatgata 920

DB 775 atccatgtc----- 783
 QY 921 aattgtcatgtcttaacacagtgctatcacagacgtcagcttgagctatgtcgaag 980
 DB 784 -----aagggacagcatgtgcttcttcaag 811
 QY 981 tgcagccgcgaagtaacgcttccgcttcccaacgctgcgcgtcgaacgtcttgcctc 1040
 DB 812 tccagcccccgaagtaacgcttccgcttcccaacgctgcgcgtcgcgtcgtcgtcgtc 871
 QY 1041 tgtatcttgcctctctgagatcagagacagacttccctccagtgatgagctgcgcgtg 1100
 DB 872 tctacccctgcgaagacagcttccccaacgctcagaatctcttccatagatctgctctg 931
 QY 1101 acggtgtcgtcttgaaggcgctgttgcacactacactctgacatctatagctgcgagc 1160
 DB 932 atgctgtctccttcaagcccgcttccagaaccttaacctctctgctgtgcgcgagc 991
 QY 1161 gctggagagtgtagtgcagcttccacacttgcgtgcgaagtcacatagatccgaacc 1220
 DB 992 gttacgagatcatattgacttaccacaacttctgtgcgagacttgcagcttgcgaacg 1051
 QY 1221 ttctgtgtgtgagctgtctggtgtgagctgtgagttgttgaacacttgacaagtgatgc 1280
 DB 1052 tgcgtgagacacagagatgcgcgacagagatgagatgcgtcagctctcagagtgatgc 1111
 QY 1281 gattcgtctgtatgagtgctcgtgagctgcgcacacttctgaggtgcgtcccaacctc 1340
 DB 1112 gctcgtgcagctcgtcgtcgtcgtgag---gacaacagcaggtccctccactctcc 1168
 QY 1341 gagaattccttcccgagggcgcaacttggagcccgcaaaccccaactgtatgacagaga 1400
 DB 1169 gtacgtctcttctcc-----tccatcaaggaaggcccgccgacgaagc 1213
 QY 1401 cttaacactcgcgcgtgtcctaagagcagtgagcaatcaacaggaatctactcttgatg 1460
 DB 1214 acttcaagtgtgaacgacagacagacacacttgcacagatgttgcgttgcgcagag 1273
 QY 1461 tgcgagacgctcgtcgcgcgaatgtgcgcgcacactgttgcagatctgcgcagctgaga 1520
 DB 1274 tcaatagcagtgctcgtgcgcgaagccgagctgcgcacgcttgcgtgcgtgcgtgcgag 1333
 QY 1521 acaactccaaaggttgcagctacacctgttcaacatctacactcgttgcagctccctt 1580
 DB 1334 acctccttgcagagcttgcagcaccgctccacatctacacttgcagcttgcagatccca 1393
 QY 1581 ctgcgtcactgcgcgtg---agtcgagccttataagtgctgtgttcaagaagctgtg 1637
 DB 1394 agcgaactgtggtcgtgcgcagagtcacagcctacagatctgcgtcttcaagaagctgcg 1453
 QY 1638 tctgcgtcgtcgtcgtgcgtgagtggtctatgttgaagcccaactacgctccttccgtaag 1697
 DB 1454 tctggttgcgcaggggtgcagaccttgcagctcagagcccaactacacactgcagc----- 1508
 QY 1698 ttctgccttcttccactaactgttctcaactcatgtcaacatcaaatggtgttcaaat 1757
 DB 1509 -----tggagcttaccat 1520
 QY 1758 gttgcacttccaaactgtatccagagagacacagacagatgagcttcaatgataatg 1817
 DB 1521 gtgcacactgcacaacacttcaacttcaagagatacgcacagatgagcttcaacgctc 1580
 QY 1818 tgtctcgttgcactagctacacacacacagatcttcaatgacacacatgagagctctgcg 1877
 DB 1581 cgcacatggagagagagatattctcaggaagagcttccgagaccatagatcccaagag 1640
 QY 1878 gaggccgcgccttccctcctcgtgagagatgcgaatgtcgcgtgcgtgactcagagact 1937
 DB 1641 gcgcgcgttcttcaacacgcagacacttccatctgcgcgtgcgtgaaacttccgcgca 1700
 QY 1938 tgcacacttgcagcagcttcaagagatgagcttcaaccccttaagccagagctgata 1997
 DB 1701 gtccatcactgcccagagtcagagagctgcgcgagacagagccgttcaacacgctcgtaga 1760


```
Dh 932 atgtgtgtctccttcaagccccgttcaagaccttaaccttaacctgtgtgtgcgagc 991
Oy 1161 gctggagagtgatgacttctccacctgtgtgcgagctcatgatatccgacc 1220
Dh 992 gttacagatcatattgacttcaacacatttgcgtgcagacttgcacctgtgcagc 1051
Oy 1221 ttctgtgtgtgcaggtctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1380
Dh 1052 ttgtgtgacacacagatgtgcgagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1111.
Oy 1281 gattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1340
Dh 1112 gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1168
Oy 1341 gaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1400
Dh 1169 gtgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1213
Oy 1401 ctctcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1460
Dh 1214 acttcaagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1273
Oy 1461 tctgagacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1520
Dh 1274 tcaatgagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1333
Oy 1521 acacatccacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1580
Dh 1334 accctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1393
Oy 1581 ctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1637
Dh 1394 agcgaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1453
Oy 1638 tctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1697
Dh 1454 tctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1508
Oy 1698 tctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1757
Dh 1509 -----tggagcttcat 1520
Oy 1758 gttgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1817
Dh 1521 gttgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1580
Oy 1818 tgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1877
Dh 1581 gttgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1640
Oy 1878 gttgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1937
Dh 1641 gttgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1700
Oy 1938 tttgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1997
Dh 1701 gttgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1760
```

```
RESULT 10
AA25727
ID AA25727 standard: cDNA: 1791 BP.
XX
AC AA25727;
XX
DT 05-JAN-2000 (first entry)
XX
DE Stachybotrys chartarum phenol oxidising enzyme encoding cDNA.
XX
KW Stachybotrys chartarum; phenol oxidising enzyme; colour; dye;
XX detergent; anti-dye transfer; stain removal; bleaching; ss.
OS Stachybotrys chartarum.
```

```
XX
PN MO9949010-AZ.
XX
PD 30-SEP-1999.
XX
PF 23-MAR-1999; 99MO-EP02042.
XX
PR 24-MAR-1998; 98US-0046969.
XX
PR 22-DEC-1998; 98US-0218702.
XX
PA (UNIL ) UNILEVER NV.
XX
PI (UNIL ) UNILEVER PLC.
XX
PI Convents D, Amory A, Wang H, Dhasee P, Lambrechts-Rongvaux A;
XX Wang C;
XX
DR WPI; 1999-601211/51.
XX
DR P-PSDB; AAY45222.
XX
PT Detergent composition containing phenol oxidase from Stachybotrys, used
XX to bleach stains and prevent dye transfer
XX
XX Example 15; Fig 5; 56pp; English.
XX
CC The present invention describes a detergent composition containing a
CC purified phenol oxidising enzyme derived from Stachybotrys. The present
CC sequence encodes Stachybotrys chartarum phenol oxidising enzyme. The
CC enzyme can be used to modify the colour of dyes and other coloured
CC compounds (e.g. for use in pulp and paper bleaching also for removing
CC stains, e.g. food, tea, blood etc., from fabrics) and for preventing dye
XX transfer during fabric washing.
XX
SQ Sequence 1791 BP; 380 A; 554 C; 448 G; 409 T; 0 other;
```

Query Match 21.9%; Score 458.8; DB 20; Length 1791;
Best Local Similarity 59.2%; Pred. No. 3,5e-119;
Matches 995; Conservative 0; Mismatches 522; Indels 163; Gaps 6;

```
Oy 321 accgtcccaaccccaacactgtgagagacatcttgaacagatgagatagacc 360
Dh 241 atattacaaacccctgtcacccggaagacattgttatcatgatatcgaacga 300
Oy 381 ttctccacagatctaacctgtatctgagcggcaacatgtgtgtgtgtgtgtgtgt 440
Dh 301 ttcaagaaagatttaacccacactgtgcctgcacactcgtgtgtgtgtgtgtgtgt 360
Oy 441 tcccaagactcatcatcgttctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 500
Dh 361 agccctgtctacttcttcaatgttccagagagacagagactgtatgtgtgtgtgtgtgtgt 420
Oy 501 agcggagagaaacacttcccaacagctcactgtcaagcgtcttctcgtagctcc 560
Dh 421 a-----atgcacacggtgaggaactcgtcatctgtcaagcgtcccaactcgtgtccct 474
Oy 561 ttgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 620
Dh 475 ttgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 534
Oy 621 aacaggagagctgcgcgaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 680
Dh 535 aactacaacatcgcgcgcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 594
Oy 681 aacgctacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 740
Dh 595 aatgcctactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 654
Oy 741 aactcccaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 800
Dh 655 gttcttccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 714
Oy 801 aacggagagcactctcttccacaatgagaggtttcagctctgtgtgtgtgtgtgtgt 860
```

```

Db 715 aacgcgaatgtaaccctcgctcgaccaggggtgagaccaggaacctggtggagatgct 774
QY attcaaatgtaagttagaccatctgatgcttcaagatcagaagtaacgtatgta 920
Db 775 atcacaatgc----- 783
QY 921 aattgtcatgctcacaacagtgctatcacagaacggtcagccttgagcctaagc 980
Db 784 -----aaagacagacatgagccttccctaagc 811
QY 981 tgcagcgcgcgaagtacagcttcgccttcctcaagctgcgcgtcctaagcttgcgtc 1040
Db 812 tccagccgcgcgaagtacagcttcgccttcctcaagctgcgcgtcctaagcttgcgtc 871
QY 1041 tgcattctgtaacctctgaagatctcagaagacagcacttcctcctaagcttgcgtc 1100
Db 872 tctaacctgcgtcagacagcagcttcctcaagcttcctaagcttgcgtcgtcgtc 931
QY 1101 acgggtgctgcttgagggccctgttgacacatctgacatctcctaagcttgcgtc 1160
Db 932 atgctgcctcctcctaagcccccgtcagaacctcaacctcctaagcttgcgtcgcagc 991
QY 1161 gcctggagagtggtatcagacttctcagccttcgctgcccagctcatalcagataccgaacc 1220
Db 992 gtcaacagacatcatctgacttcaacaacttgcgtccagactcttgacatcgcgaacg 1051
QY 1221 ttcctgtgctgacagcttcctcgctgctgagccttgatctgaacaactgacaagctacgc 1280
Db 1052 tctgtcagacacacagatgctgcgacagagatgagtaagctgcacactcctgaagctatgc 1111
QY 1281 gattcgtcgttgatgaagctccttgagtcgcgcgaacctctgaggtgctgcgaacctcc 1340
Db 1112 gctcgtcgtcagctcgtgacatgttgag--gacaacagcagatgcctcctaactcctcc 1168
QY 1341 gaagatgcttccttcctccgaagggcgcgcaactgtagccccaaccccaactgagtagaga 1400
Db 1169 gtagagcttccttccc-----tcccaacaagaagagccgcgcgaacagc 1213
QY 1401 ctctcaccttcggcgcgtgtaataagacagtgacaactcaacagagatctcctcgatg 1460
Db 1214 actcaagtttgagacgagacagacgacactaactgatacagatgctgtgttcgcagatg 1273
QY 1461 tgcagaacacgtctgcctcgcgaatgtgcctccgcgacactgttgatctgagcagcttgaga 1520
Db 1274 tcaataagcgtctgcctcgcgaacgacccgagcctgcgcacgcgttgagcttgaggaactcgaga 1333
QY 1521 acaacacacacaggttgagacacacacacacacacacacacacacacacacacacacacac 1580
Db 1334 acctcctcgtgagcgtgagacacacacacacacacacacacacacacacacacacacacac 1393
QY 1581 ctgcgtcacaatgcgcgttg--agtcgagcctatgagcgtgctgctcctaagatgttg 1637
Db 1394 agcgaactgtgtgctgctgagcaggtcagcctaagctgctgtcttaagatgtgcg 1453
QY 1638 tctggtgctgctgctgctgaggtgtgtatgtgtgagcccaactacgcttccctccgtaag 1697
Db 1454 tctggtgctgagcaggtgtgagacacacacacacacacacacacacacacacacacacacac 1508
QY 1698 tctgcgccttcaactaactgttctcaatcatgatacaatacaacaagtgtgtctcaac 1757
Db 1509 -----tggaactcaat 1520
QY 1758 gttgcaatcgcacaacatctacacagagacacacacacacacacacacacacacacacacac 1817
Db 1521 gttgcaatcgcacaacatctacacagagacacacacacacacacacacacacacacacacac 1580
QY 1818 tgttctcgttgatgctgtaacacacacacacacacacacacacacacacacacacacacac 1877
Db 1581 cgcacagagagagagagatcttcaagagagacacacacacacacacacacacacacacacacac 1640
QY 1878 gaagcccccgcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1937
Db 1641 gcgcgcgcgttctccttaacacacacacacacacacacacacacacacacacacacacacac 1700

```

```

QY 1938 tgcatactgacgcacatcagagagatgcttcaacccctacagcccaagcgtatga 1997
Db 1701 gtccatacactgcccagagtgagagagctgagcagcagagacgtagacacacacacacac 1760

```

```

RESULT 11
AA25735
ID AA25735 standard; DNA; 2067 BP.
XX
AC AA25735;
XX
DT 05-JAN-2000 (first entry)
XX
DE Stachybotrys chartarum phenol oxidising enzyme PCR fragment.
XX
KW Stachybotrys chartarum; phenol oxidising enzyme; colour; dye;
KW detergent; anti-dye transfer; stain removal; bleaching; PCR primer; ss.
XX
OS Synthetic.
XX
SS Stachybotrys chartarum.
XX
PN MO949010-A2.
XX
PD 30-SEP-1999.
XX
PF 23-MAR-1999; 99WC-EP02042.
XX
PR 24-MAR-1998; 98US-0046969.
PR 22-DEC-1998; 98US-0218702.
XX
PA (UNIT ) UNILEVER NV.
PA (UNIT ) UNILEVER PLC.
XX
PI Convents D, Amory A, Wang H, Dhaese P, Lambrechts-Rongvaux A.
PI Wang C;
XX
DR WPI; 1999-601211/51.
XX
PT Detergent composition containing phenol oxidase from Stachybotrys, used
PT to bleach stains and prevent dye transfer -
XX
PS Example 17; Fig 9; 56bp; English.
XX
CC The present invention describes a detergent composition containing a
CC purified phenol oxidising enzyme derived from Stachybotrys. The present
CC sequence represents a PCR fragment of Stachybotrys chartarum phenol
CC oxidising enzyme. The enzyme can be used to modify the colour of dyes
CC and other coloured compounds (e.g. for use in pulp and paper bleaching
CC also for removing stains, e.g. food, tea, blood etc., from fabrics) and
CC for preventing dye transfer during fabric washing.
XX
SQ Sequence 2067 BP; 445 A; 621 C; 500 G; 501 T; 0 other;

```

```

Query Match 14.3%; Score 300.4; DB 20; Length 2067;
Best local Similarity 55.2%; Pred. No. 2,4e-74;
Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;

QY 380 ctctccacaacatctacacatctgacatcgcgcgcacacatggttgatagatgacat 439
Db 457 ctgaccccttcagatthaccacacacacacacacacacacacacacacacacacacacac 516
QY 440 gtcccaagacacacacacacacacacacacacacacacacacacacacacacacacacac 499
Db 517 gagcctgtctcacttcaatgattccacagagagacacacacacacacacacacacacacac 576
QY 500 cagcagagagacacacacacacacacacacacacacacacacacacacacacacacacac 559
Db 577 ca-----atgcacacgttgagacacacacacacacacacacacacacacacacacacac 630
QY 560 ctctgattgttggtcgtgagacacacacacacacacacacacacacacacacacacacacac 619

```


RESULT 14
AAZ27602
ID AAZ27602 standard: DNA; 3677 BP.
XX
AC AAZ27602;
XX
DT 16-DEC-1999 (first entry)
XX
DE Stachybotrys phenol oxidase genomic sequence.
XX
KM Phenol oxidase; enzyme; coloured compound; dye transfer prevention;
KW fabric washing; stain bleaching; anti-dye transfer; detergent; ss.
XX
OS Stachybotrys chartarum.
XX
PN W09949020-A2.
XX
PD 30-SEP-1999.
XX
PE 23-MAR-1999; 99WO-US06327.
XX
PR 24-MAR-1998; 98US-0046969.
PR 22-DEC-1998; 98US-0218702.
PR 22-MAR-1999; 99US-0273957.
XX
PA (GENEV) GENENCOR INT INC.
XX
PI Amory A, Wang H, Dhase P, Lambrechts-Rongvaux A, Wang C;
XX
DR WPI: 1999-591088/50.
DR P-PSDB: AAY39992.
XX
PT Novel enzyme for modifying coloured compounds used to prevent
XX
PS dye-transfer -
XX
PS Claim 21; Fig 6; 64pp: English.
XX
CC This sequence encodes the Stachybotrys chartarum phenol oxidase enzyme
CC of the invention. The invention is used to modify a coloured compound and
CC prevent dye transfer during fabric washing, or for stain bleaching or
CC anti-dye transfer. It is useful in the detergent, paper and pulp, textile
CC and food industries.
XX
CC Sequence 3677 BP; 822 A; 1056 C; 849 G; 947 T; 3 other:
SQ

Query Match 14.3%; Score 300.4; DB 20; Length 3677;
Best Local Similarity 55.2%; Pred. No. 3.3e-74;
Matches 957; Conservative 0; Mismatches 501; Indels 276; Gaps 8;

QY 380 cttctccaccagatcatctgactgtgagcgcgccaacatggttgatagatgcat 439
DB 1491 cgcacccttcagatctaccacacacttcgcgcctgcaactctgcgtacagatggat 1550
QY 440 gtcccaagaccatcatcatcgttcctcgttgcaatgagatgtgtccgcttgaa 499
DB 1551 gggccttgctcactctcaatgttccagagaggaacagagactgtatgtatgctaa 1610
QY 500 cagcgagagaaacactctcccaacagcgtccaacttgacagcgtctttctcagctcc 559
DB 1611 ca-----atgcacacgttgagaaactcgttcctcgtccacgtcccatcgctgcc 1664
QY 560 ctttgatgttggttgagagacacacacacgttgagagatgaagattactaaccc 619
DB 1665 tttcgtatgttggttgagagatgtgacctctccctcgttgagatgaagattactaac 1724
QY 620 caacaagcagcgttcgcagatcgttgatcacaatgacatgc-----661
DB 1725 caactacacatcgcgcgcctctcgtgtacacatgacacgctcttaatgaagtatgcta 1784
QY 662 -----catgtccatcacgcgcag 680

DB 1785 cgagccttactcttcttgctacacttcttgccaaccacttctctgtagactgctgag 1844
QY 681 aacgcctacatgttgtaagcctgtgtctctacatgatccaagcgcgttagatgctccg 740
DB 1845 aatgcctactcttgtaagcctgttgccaactatatacaacagagcgtgagatgctccc 1904
QY 741 aactccccaagcgcgttgtaagcctgttgccaactatatacaacagagcgtgagatgctccc 800
DB 1905 ggtcttcctcagtgagcgttgtaagcctgttgccaactatatacaacagagcgtgagatgctccc 1964
QY 801 aacgcagagcgcgtctctctccacaaatgagaggtttccagctctcgtggtgagctt 860
DB 1965 aacgcagagcgcgtctctctccacaaatgagaggtttccagctctcgtggtgagctt 2024
QY 861 attcaagtggtgaggttgagcccatgtgagctgtctcagacccagagatgcgatgata 920
DB 2025 atccatgtc-----2033
QY 921 aattgtgatgtctctaacacagtgctatacagaacggtcagcttggtcctatgcaag 980
DB 2034 -----aacgcagacagcatgctcttccctaaag 2061
QY 981 tgcagcgcgcgaagtaacgccttcgccttcctcaacgctgcgtctcaagctcttgcctc 1040
DB 2062 tccagcccgcaagtaacgccttcgccttcctcaacgctgcgtctcaagctcttgcctc 2121
QY 1041 tgatcttctcactctctgaggttcagagacagactctctccagcagctgctgctgctg 1100
DB 2122 tctaacctcgtcagagacagctccccaacgctcagaatctcttcccaagcatgtgctctg 2181
QY 1101 acggtgtctgtctgagagcctcgttgtaacatgacactctgtacatctataagcagag 1160
DB 2182 atgctgtctccttcaagcccgctcagacactcctacactcctgctgtgctgagcagc 2241
QY 1161 gctgagaggttgta-----1175
DB 2242 gttacgagatcatattgtgtatgcccctccctccctcacaagatgataaagactcaag 2301
QY 1176 -----tcgaacttcacactcgtcgtgagcagtcacatcagatccgaacactcctg 1226
DB 2302 taacactttagactcaccacaaacttgcgtgcagactcttgacgtgcgaacgttgctg 2361
QY 1227 gttgtacggtctcgtgtgttgagcctgagttgtataacactgacagatgacgtatcg 1286
DB 2362 agaccaagatgctgcgcgcagagatgagtagctgcgactcgcaggtgagatgctgctg 2421
QY 1287 tggttatgaagtctctgtgagtcgcgcgaacactcttgaggtgctgcgaactccgagatg 1346
DB 2422 tctgcagctctgtgcaactgttgag--gacaacagccaggtlccctcctcactctcgtgagc 2478
QY 1347 ttccttcccgagagggcgcaacttgagcccgcaaacccaactgtagtcgagacttca 1406
DB 2479 ttccttccct-----cctcaacaggaagggcccgcgcaagcaacttca 2523
QY 1407 ccttcgcccgttgtaatgacagtgtagacaatcaacaggaattacactctcgtatgtcaga 1466
DB 2524 agtttgaacgcagcaagcaactacacttgatcaagatgttgcttgccgtatgcaatg 2583
QY 1467 accgtctgtctcgcgaatgtgccccgcgcagacagttgagatctcgggcacttgagaacact 1526
DB 2584 agcgtgtctctgcaagcccgagctgagcagctgttgaagctctggagcctcgagaactct 2643
QY 1527 ccaaggttgtagactcacccgtgtacatcaactcgttgacttcgcagctccttcgctt 1586
DB 2644 ctggagagctggagcaaccccgctccacatcaactctgttacttaagaatccctcaagcgaa 2703
QY 1587 ccaactgccgttg--agtcgaagccttatagagctgtgctcacaagatgtgtctcagc 1643
DB 2704 ctggtgtgtgtgagcagatgcatagcctcagagctgtcgtcttaagagatgtcgtctgt 2763
QY 1644 tggctcgtctgagaggtgtgtctatgttgagagcccaactagctccttcccgtaagttctcg 1703
DB 2764 tgggcaggggtgagacccttgacatcagagcgacactacacacacttgagc-----2812

Yr	Accession	Sequence	Position
0y	1704	cctttactaacaggttttactcatgtataactaanaagttgtgtctacagttga	1763
Db	2813	-----tggagcttacaatgtgca	2830
0y	1764	ctgcacaacctgataccagagagacacagacatgatgctgtcttcaatgtcacgttct	1823
Db	2831	cttcaacacactcattacagagataagacatgatagtgtatttcaacgtcacgcgcac	2890
0y	1824	cgtgactactatggtctacacaactcaacccagatctatgaccccaatggagctctctggagcc	1883
Db	2891	ggagagaaagggaatcttcaagagagacttcagagaccaccaatgaaccccaagttgctgagc	2950
0y	1884	ccgcccctctccctctcgagagatctcagagatgctcgtggatcttcacgagcttgcac	1943
Db	2951	cgttccttacaacgcgaacgacttcacatgctcgtcgctgtaaaccttctccgcgcgtccat	3010
0y	1944	caatgacgcgcatcgaagagatggtctgacttcaacccctacgcgccagctgtga	1997
Db	3011	caatgcgccgagttgaagagagctgcccgaagagaagccgtacacacgcctcagatga	3064

RESULT	15
AAAS0018	
ID	AAAS0018 standard; DNA; 3677 BP.
XX	
AC	AAAS0018;
XX	
DT	10-OCT-2000 (first entry)
XX	
DE	Stachybotrys chartarum phenol oxidising enzyme gene.
XX	
KW	Phenol oxidising enzyme; detergent; bleaching; ds.
XX	
OS	Stachybotrys chartarum.
XX	
EH	Key
FT	Location/Qualifiers
FT	CDS
FT	1044..3095
FT	/*tag= a
FT	/note= "contains introns"
FT	1044..1237
FT	/*tag= b
FT	/number= 1
FT	1238..1285
FT	/*tag= c
FT	1286..1324
FT	/*tag= d
FT	/number= 2
FT	1325..1376
FT	/*tag= e
FT	1377..1449
FT	/*tag= f
FT	/number= 3
FT	1450..1503
FT	/*tag= g
FT	1504..1776
FT	/*tag= h
FT	/number= 4
FT	1777..1835
FT	/*tag= i
FT	1836..3095
FT	/*tag= j
XX	
PN	MO200039306-A2.
XX	
PD	06-JUL-2000.
XX	
PF	20-DEC-1999; 99WO-EP10287.
XX	
PR	23-DEC-1998; 98US-0220871.
PR	23-JUN-1999; 99US-0338723.
XX	
PA	(UNIL) UNILEVER NV.

PA (UNIL.) UNILEVER PLC.
PA (HIND-) HINDUSTAN LEVER LTD.
XX
PI Bodie EA, Van Der Velden S, De Vries CH, Wang H
XX
DR WPI: 2000-514528/46.
DR P-PSDE: AAY95357.

WPI; 2000-514528/46.
P-PSDB; AAY95537.

Detergent composition comprising novel phenol oxidising enzyme obtained from fungus or bacteria, useful for pulp and paper bleaching, bleaching color of stains on fabric and for anti-dye redeposition -

Claim 1; Fig 1A-B; 45pp; English.

The present sequence is that of the *Stachybotrys chartarum* MUC1 38898 phenol oxidising enzyme gene, including promoter and terminator sequences. The gene was isolated from genomic DNA using primers (see AA50023-24), based on isolated peptides of the enzyme. The gene codes for a 594-amino acid protein (see AA195537). The invention relates to detergent compositions comprising novel phenol oxidising enzymes that are encoded by nucleic acids capable of hybridising to the present DNA sequence, provided the enzymes are capable of modifying the colour associated with dyes or coloured compounds, and are produced from a bacterium, yeast or fungus (see AA195538-40). The phenol oxidising enzymes can be used for pulp and paper bleaching, for bleaching the colour of stains on fabric and for anti-dye transfer in detergent and textile applications. They may also be capable of modifying the colour in the absence or presence of an enhancer. Expression vectors and host cells comprising a nucleic acid encoding a phenol oxidising enzyme, methods for producing the phenol oxidising enzyme, and methods for constructing expression hosts are provided.

Sequence 3677 BP; 822 A; 1057 C; 849 G; 948 T; 1 other;

Query Match	14.38;	Score 300.4;	DB 21;	Length 3677;
Best Local Similarity	55.28;	Pred. No. 3.3e-74;		
Matches 957; Conservative	0;	Mismatches 501;	Indels 276;	Gaps 8

[illegible]

Db 1965 aacgcagatgtgtaacccttcgctcgaacagaggtgaggaacgaacctgagggaatgttc 2024
 Qy 861 attcaagtgttaagtgtgagcccatgtgagatgtcttaagatccatagaagtatcatgtatga 920
 Db 2025 atcaatgttc----- 2033
 Qy 921 aatgtcatgtcttaaccagttgatacaagaacgttaagccttggcctatgtccaag 980
 Db 2034 -----aacggaacgacatgtccttccaaag 2061
 Qy 981 tgcagcgcgcgaagtaacgccttcgccttccttaacgcttgccgtcccaagcttcttcgcgc 1040
 Db 2062 tccagcccccgcgaagtaacgccttccgaatcccaacgctgcgtcgtctgtcttgctcc 2121
 Qy 1041 tgaatcttgtaacccctcgtgagatcagaggaacgaacttccctccagagcatgtccgctg 1100
 Db 2122 tctacctcgtcaggaacagctcctccaaagctagaatccttccaaagcatltgtccttg 2181
 Qy 1101 acggtgtctgtcttgagggccctgtgtgaacactgaacactgtacalctcatagccgagc 1160
 Db 2182 atgctgtctctcttaagcccccgttcaagacttaaccttaacctgtgtgttgccgagc 2241
 Qy 1161 gctggaggtgtgtta----- 1175
 Db 2242 gttacagagatcatltgtgtatgcctccctcctcaagaatgagtaagaactctaagac 2301
 Qy 1176 -----tgacttcccaactcgcctcgcagctccatccatccgaactccgcaacttcctg 1226
 Db 2302 taacactgttaagacttaacaaacttgcgtgcgaacactctgaactcgtgcgaacgltgcg 2361
 Qy 1227 gtcgtgaacgtctcgtgtgttgagcctgagttgtataacaactgaacaggtcatgtcga 1286
 Db 2362 agaccaacagtctgcgcgcagagatgagtaagctgcagactcctgcaggtgatgcgttcg 2421
 Qy 1287 tccgttgatgaagtctccttgagatgcgcgcgaacactctgaggtgtcctgcgaactccgagatg 1346
 Db 2422 tccgtgaactctgcagactgttgag--gacaacagccaggtccctccactctcgttgagc 2478
 Qy 1347 ttccttcccccaggggcgaacactggaacccgcgaaccccaactgaatgaaggaacttca 1406
 Db 2479 ttccttccct-----cctcaagaagaagcccccgcgcgaacgaacacttca 2523
 Qy 1407 ccttcgcgcgtgtcctaagtgaacagtgaaacaaatcaacggaatcaactctcgaatgtcgaag 1466
 Db 2524 agttgaacgcgcgaacgaacactaactgaatcaacatgtgtgttgccgaatgtcaatg 2583
 Qy 1467 accgtctgtctccgcaatgtgcgcgcgcgaacactgttgagatctgcgcgaacttgagaacaa 1526
 Db 2584 agcgtgtcctcgcgaacccgcgaacactgtgaagctcgcgcgcgaacactcct 2643
 Qy 1527 ccaacggttggaactaacctgttcaactcaactcgttgaactcgcgaactccttccgt 1586
 Db 2644 ctggaggtcgcgcgcgaacccgcgcgaacactcaactcgttgaactcaagatccccaagcgaa 2703
 Qy 1587 ccaactgcgcgtg---agtcgaagccttaltgaggtgcgtgtcctcaagagatgtgtcgcgc 1643
 Db 2704 ctgggtgtcgttgccaggtcaggtccctacgagctcgtgtcttaagagatgtcgtctggt 2763
 Qy 1644 tggcctcgtcgttgaggtgtgtctatgttgagggcccaactcgtccttcccgtaagtctcg 1703
 Db 2764 tgggcaggggtgagacccctgcgaacactgaagggcccaactcaacaccctggac----- 2812
 Qy 1704 ccttlaacttaactgttttcaactatgtcaactaactaagaatgtgttcaatagtgtga 1763
 Db 2813 -----tgagacttaactgtgca 2830
 Qy 1764 ctgcacaacactgtatccagaggaacgaacgaatgtatgtcgtctcaatgtcaatgtctc 1823
 Db 2831 ctgtcaacaacactcaactgaagatgaacgaacatgtatgtctatcaacgtcaccccat 2890
 Qy 1824 cgtgtactatgtctataactaaccgaagttcaatgcccatggagcctctcttgagagcc 1883

Db 2891 ggaagagaagagatatcttcaaggaagacttcgaggaaccccaatgaaccccaagtgtgcgc 2950
 Qy 1884 ccgccccttccctcctcctcgcgcgaaggttcgagaaatggtcgcgtggtgaactcaacgagcttgcac 1943
 Db 2951 cgttcccttaacacgaacgaacttccatgtcgtcgtgaaacttccgcgcgcgagtcacat 3010
 Qy 1944 caactgaacgcatctcagagagatgtctagcttaacacccctacgcccagcgtgata 1997
 Db 3011 cactgcgcaggtgcagaggtgtgcgcgaacgaagagccgtacaaacgcctcgtatga 3064

Search completed: October 5, 2002, 20:29:54
 Job time: 9023 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 18:54:14 ; Search time 3211.21 Seconds

(without alignments)
8805.441 Million cell updates/sec

Title: US-09-656-640A-3

Perfect score: 2095

Sequence: 1 cagctcggtctactctc.....gaataagttgtgtcttaa 2095

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estlum:*
3: em_estlin:*
4: em_estmu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estfro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86.4	4.1	600	BE188240	M7ATIG In
2	86.4	4.1	700	BE187716	BTMOX In
3	68.6	3.3	664	BE188099	CRC377-R
4	65.6	3.1	848	BE216983	EST0376-T
5	63.4	3.0	880	BE264805	HV_CEA001
6	54.6	2.6	465	BC053739	BC053739_RHR2.9_B
7	51.4	2.5	311	AV426401	AV426401_NY426401
8	50	2.4	580	BE159802	BE159802
9	48.2	2.3	592	BE143324	BE143324
10	47.8	2.3	766	BC526127	56-88 Ste
11	46.2	2.2	286	A2714471	RPCR-24-1
12	46.2	2.2	599	BE258534	HVMEF001
13	46.2	2.2	669	BE258534	HVMEF001
14	45.6	2.2	485	BE1876848	UJ35B06.Y
15	45.6	2.2	562	BE651243	BB651243
16	45.6	2.2	596	BE1574420	uk20h11.Y
17	45.6	2.2	600	BE634973	BE634973

18	45.6	2.2	634	9	AI574382	AI574382 uk20h11.Y
19	45.6	2.2	738	10	BE918653	BE918653 602819588
20	45.6	2.2	762	9	AI316242	AI316242 uJ26F06.Y
21	44.8	2.1	657	10	BE1527303	BE1527303 1024080A0
22	44.2	2.1	705	12	BE590982	BE590982 BOCWH85TR
23	43.8	2.1	433	9	AW018870	AW018870 fd61D03.Y
24	43	2.1	552	12	A2302933	A2302933 GSSBRu187
25	42.8	2.0	473	9	AV633110	AV633110 AV633110
26	42.8	2.0	542	9	AV626368	AV626368 AV626368
27	42.8	2.0	611	9	AV391222	AV391222 AV391222
28	42.6	2.0	725	9	AI317680	AI317680 uJ22C09.Y
29	42.4	2.0	533	9	BE234675	BE234675 141860 MA
30	42.4	2.0	551	10	BE321929	BE321929 uz66d04.Y
31	42.4	2.0	591	10	BE789685	BE789685 602103709
32	42.4	2.0	697	10	BE914800	BE914800 602813607
33	42.4	2.0	795	10	BE1559124	BE1559124 603241206
34	42.4	2.0	885	10	BE1558920	BE1558920 603240570
35	41.8	2.0	667	9	BE212100	BE212100 894037F06
36	41.6	2.0	444	10	BM100862	BM100862 EBP101.S0
37	41.6	2.0	518	9	AA182405	AA182405 zp36q07.r
38	41.2	2.0	339	10	F01214	F01214 HSBSE082.S
39	41.2	2.0	541	10	BE255913	BE255913 601109883
40	41.2	2.0	722	10	BE250985	BE250985 601107071
41	41.2	2.0	828	10	BE749472	BE749472 602707601
42	41.2	2.0	1128	12	AF094955	AF094955 AF094955
43	40.8	1.9	518	9	BE031325	BE031325 129855 MA
44	40.8	1.9	534	10	BM421598	BM421598 V017A09.O
45	40.6	1.9	316	10	BE725648	BE725648 894085A09

ALIGNMENTS

RESULT 1
LOCUS BE188240 600 bp mRNA linear EST 25-SEP-2000
DEFINITION M7ATIG In vitro expressed cDNAs Cladosporium fulvum similar to
Bilirubin oxidase precursor. mRNA sequence.
ACCESSION BE188240
VERSION BE188240.1 GI:8667479
KEYWORDS
SOURCE Cladosporium fulvum.
ORGANISM Cladosporium fulvum
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothyrionmycetes Incertae sedis; Mycosphaerellaceae; mitosporic
Mycosphaerellaceae; Cladosporium.
REFERENCE 1 (bases 1 to 600)
AUTHORS Clark,A.J., Rasmussen,S.W. and Oliver,R.P.
TITLE In vitro expressed genes of Cladosporium fulvum
JOURNAL Unpublished (2000)
COMMENT Contact: R.P.Oliver
Neurotrophic Phytopathology Research Centre
Murdoch University
SABC, Perth 6150, Western Australia
Tel: +61-8-9360-7404
Fax: +61-8-9360-6303
Email: roliver@central.murdoch.edu.au
High quality sequence stop: 600
POLYA-No.

FEATURES

source
1..600
/organism="Cladosporium fulvum"
/strain="Race 4"
/db_xref="taxon:5499"
/clone_lib="In vitro expressed cDNAs"
/tissue_type="Mycelium"
/note="Vector: lambda bluescript; A mixture of mycelial cultures grown in liquid B5 for 48 hours and transferred for 24 hours to media lacking carbon, nitrogen or supplemented with hydrogen peroxide"
BASE COUNT 131 a 176 c 128 t 3 others
ORIGIN

Query Match 4.1%; Score 86.4; DB 9; Length 600;
Best Local Similarity 55.0%; Pred. No. 4.8e-11;
Matches 243; Conservative 0; Mismatches 178; Indels 21; Gaps 3;

```
OY 1253 gagttgataaacactgacaaagatcatgcatctgcttgatgaagtccttgatgcgc 1312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 64 GAGTGCATTAACACCGACACAGGTCATGAGTGTGCTGGAGA-----C 108
OY 1313 gacactctgagtgctcctgcaacctcgaagatgtctcttcccgagggcggaactgg 1372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 109 AGCGTACCGACGAGTGCACAAACAGTACCATCACTTAACGCGTGCATCGACTGG 168
OY 1373 gaccccgaaaccccaactgatacgaagacttcaactc---ggccgtgctaatagacag 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 169 CCAGCGCAGACGACACTATTGACAAGACCTTCAACTCCAGATGGCGCGGACGAGTG 228
OY 1430 tggacaataacagaggttcccttcctgcatgtgcgagaacggtctgctcgcaatgtgcgc 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 TGGACCGTCAATGGCGTGGATTTCAGCGACCCGAACTCGCTGTGCTGAGGCCGCA 288
OY 1490 cgcgaacactgtgagatcctgagacttgagaacaaactcaacggttgagactcaacctgt 1549
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 288 CAAGGACACTGTGAGAGGTTGGCGCTCGTGCACACTGGGGTCCGGCTGTCCACCCAGTA 348
OY 1550 caactacactcgttgacttcgagtccttccttcgttc---cactgcccgtgagtcgag 1606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 349 CATATCCATCTTGTCAACATGACAGTCCCTCGCGTACTGTGTGTCGCCGTGCTCATG 408
OY 1607 ccttctgagcgtcgtgctcgaagatgtgtctgctgctgctgctgctgagtgctat 1666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 409 CCATACGAGACTGCTGTGTGTAAGACGTCGTATGCTGTGCTGCGAGATTGTGAC 468
OY 1667 gttgagcccaactacgctcct 1688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 469 GTGCTTGCTTTACGGTCTT 490
```

RESULT 2
BE187716 700 bp mRNA linear EST 25-SEP-2000
LOCUS BE187716
DEFINITION BILIOX In vitro expressed cDNAs Cladosporium fulvum cDNA similar to
Bilirubin oxidase precursor, mRNA sequence.
ACCESSION BE187716
VERSION BE187716.1 GI:8666955
KEYWORDS EST.
SOURCE Cladosporium fulvum.
ORGANISM Cladosporium fulvum
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothyriomycetes incertae sedis; Mycosphaerellaceae; mitosporic
Mycosphaerellaceae; Cladosporium.
REFERENCE 1 (bases 1 to 700)
AUTHORS Clark,A.J., Rasmussen,S.W. and Oliver,R.P.
TITLE In vitro expressed genes of Cladosporium fulvum
JOURNAL Unpublished (2000)
COMMENT Contact: R.P.Oliver
Neurotrophic Phytopathology Research Centre
Murdoch University
SABC, Perth 6150, Western Australia
Tel: +61-8-9360-7404
Fax: +61-8-9360-6303
Email: roliver@central.murdoch.edu.au
High quality sequence stop: 700
POLYA-No.

FEATURES
source Location/Qualifiers
1. 700
/organism="Cladosporium fulvum"
/strain="Race 4"
/db_xref="taxon:5499"
/clone_lib="In vitro expressed cDNAs"
/tissue_type="Mycelium"
/note="Vector: Lambda bluescript; A mixture of mycelial
cultures grown in liquid B5 for 48 hours and transferred."

for 24 hours to media lacking carbon, nitrogen or
supplemented with hydrogen peroxide 7 others
BASE COUNT 158 a 206 c 183 g 146 t
ORIGIN

Query Match 4.1%; Score 86.4; DB 9; Length 700;
Best Local Similarity 55.0%; Pred. No. 5.3e-11;
Matches 243; Conservative 0; Mismatches 178; Indels 21; Gaps 3;

```
OY 1253 gagttgataaacactgacaaagatcatgcatctgcttgatgaagtccttgatgcgc 1312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 64 GAGTGCATTAACACCGACACAGGTCATGAGTGTGCTGGAGA-----C 108
OY 1313 gacactctgagtgctcctgcaacctcgaagatgtctcttcccgagggcggaactgg 1372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 109 AGCGTACCGACGAGTGCACAAACAGTACCATCACTTAACGCGTGCATCGACTGG 168
OY 1373 gaccccgaaaccccaactgatacgaagacttcaactc---ggccgtgctaatagacag 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 169 CCAGCGCAGACGACACTATTGACAAGACCTTCAACTCCAGATGGCGCGGACGAGTG 228
OY 1430 tggacaataacagaggttcccttcctgcatgtgcgagaacggtctgctcgcaatgtgcgc 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 TGGACCGTCAATGGCGTGGATTTCAGCGACCCGAACTCGCTGTGCTGAGGCCGCA 288
OY 1490 cgcgaacactgtgagatcctgagacttgagaacaaactcaacggttgagactcaacctgt 1549
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 289 CAAGGACACTGTGAGAGGTTGGCGCTCGTGCACACTGGGGTCCGGCTGCTCACCCAGTA 348
OY 1550 caactacactcgttgacttcgagtccttccttcgttc---cactgcccgtgagtcgag 1606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 349 CATATCCATCTTGTCAACATGACAGTCCCTCGCGTACTGTGTGTCGCCGTGCTCATG 408
OY 1607 ccttctgagcgtcgtgctcgaagatgtgtctgctgctgctgctgctgagtgctat 1666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 409 CCATACGAGACTGCTGTGTGTAAGACGTCGTATGCTGTGCTGCGAGATTGTGAC 468
OY 1667 gttgagcccaactacgctcct 1688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 469 GTGCTTGCTTTACGGTCTT 490
```

RESULT 3
BE188099 664 bp mRNA linear EST 25-SEP-2000
LOCUS BE188099
DEFINITION CFC377-R In vitro expressed cDNAs Cladosporium fulvum cDNA similar
to Bilirubin oxidase precursor, mRNA sequence.
ACCESSION BE188099
VERSION BE188099.1 GI:8667338
KEYWORDS EST.
SOURCE Cladosporium fulvum.
ORGANISM Cladosporium fulvum
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothyriomycetes incertae sedis; Mycosphaerellaceae; mitosporic
Mycosphaerellaceae; Cladosporium.

REFERENCE 1 (bases 1 to 664)
AUTHORS Clark,A.J., Rasmussen,S.W. and Oliver,R.P.
TITLE In vitro expressed genes of Cladosporium fulvum
JOURNAL Unpublished (2000)
COMMENT Contact: R.P.Oliver
Neurotrophic Phytopathology Research Centre
Murdoch University
SABC, Perth 6150, Western Australia
Tel: +61-8-9360-7404
Fax: +61-8-9360-6303
Email: roliver@central.murdoch.edu.au
High quality sequence stop: 664
POLYA-No.

FEATURES
source Location/Qualifiers
1. 664
/organism="Cladosporium fulvum"
/strain="Race 4"

/db_xref="taxon:5499"
/clone_lib="In vitro expressed cDNAs"
/tissue_type="Mycelium"
/note="Vector: lambda bluescript; A mixture of mycelial cultures grown in liquid B5 for 48 hours and transferred for 24 hours to media lacking carbon, nitrogen or supplemented with hydrogen peroxide"

BASE COUNT 149 a 174 c 174 g 140 t 27 others
ORIGIN

Query Match 3.3%; Score 68.6; DB 9; Length 664;
Best Local Similarity 58.1%; Pred. No. 1.6e-06;
Matches 158; Conservative 0; Mismatches 107; Indels 7; Gaps 2;

QY 952 gaagcgtgagcgttcctgctgaagctgagcgcgcaagctaccgcttcct 1011
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 57 GAACCTGCAGCCATGCGCATCTGACATCGACCTCGACATCGACTCTT 116
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1012 caagctgcgtctcagcgtcttcctgtatctgtacctctgagatcagagac 1071
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 117 TGTATATGTCCTTACAGCAGATCTTACGATCTGATG-----CTGATGAGACACAA 170
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1072 cagacttccctccagatctatgacgtgacgtgtgtctgtgagccctgtgaac 1131
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 171 CCCAATCCATTCAGAGCTATGCTCAGACCTCCGCTATTCGCCCTCGACGC 230
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1132 tgacactctgatactctatgacgcgtgaggtgtgtatgaccttccacct 1191
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 231 CAAGAGCTGTCATTTCCATGAGAGAGCGCT-TGATCATCATTCGACTTANCTGT 289
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1192 cgtgagccagctcagatcagatccgacacctc 1223
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 290 CCGCGGCCAGATATATCTATGTTTANCGGTC 321
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
LOCUS BE216983 848 bp mRNA linear EST 03-JUL-2000
DEFINITION EST0376 Triticum aestivum Lambda Zap Triticum aestivum cDNA clone
JAI_5A_A06_T3 5', mRNA sequence.
ACCESSION BE216983
VERSION BE216983.1 GI:8904523
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Triticum.
1 (bases 1 to 848)
REFERENCE 1 (bases 1 to 848)
AUTHORS Anderson,J.M., Williams,C.E. and Goodwin,S.B.
TITLE Analysis of an EST database reveals a probable Cf2 resistance gene
homolog in wheat
JOURNAL Unpublished (2000)
COMMENT Crop Production & Pest Control Research Unit
Contact: Anderson, J.M.
USDA-ARS
1150 Lilly Hall, West Lafayette, IN 47907, USA
Tel: 765-494-5565
Fax: 765-496-2926
Email: janderson@purdue.edu
Seq primer: T3
High quality sequence stop: 848.
Location/Qualifiers
1..848
/organism="Triticum aestivum"
/strain="P29"
/db_xref="taxon:4565"
/clone_lib="JAI_5A_A06_T3"
/clone_lib="Triticum aestivum Lambda Zap"
/tissue_type="leaf"
/dev_stage="9 day old seedlings"
BASE COUNT 139 a 312 c 263 g 134 t

ORIGIN

Query Match 3.1%; Score 65.6; DB 9; Length 848;
Best Local Similarity 55.2%; Pred. No. 1.1e-05;
Matches 128; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 590 gctctgagctaacgaatctactaccacacagcagcgtccgcatgttgta 649
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 GCACGAGACTCGCCGCGCATGACGATCCGAGACGACCTCCCGATACCTCTGTGA 60
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 650 ccatgaccatgcatatgcatcaccgcgcgagacgctcatatggtcaggtgtgta 709
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 CCACGACCAAGCGCATGGGCTCACCCGCGTCAATCTCTGCTGCGCGCGCA 120
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 710 catgattccagacccgctgagatgcttgaacctcccaagaggttaaggagttga 769
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 CCGTGTGCGCGACCGCGACCGCGCGCCCTCTCCGCGCGGAGTTCA 180
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 770 talcccttggtctctacgtccagacgatacacgacgacgacgtctctc 821
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 CCGGACCTCTCTCTCTCTCTGACCGCGACTTCAGAGACGACGCGCGCTCTTC 232
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
LOCUS BF264805 880 bp mRNA linear EST 23-OCT-2001
DEFINITION HV.CE0010619f Hordeum vulgare seedling green leaf EST library
HVCNMA0004 (Blumeria challenged) Hordeum vulgare cDNA clone
HV.CE0010619f, mRNA sequence.
ACCESSION BF264805
VERSION BF264805.2 GI:13261741
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 880)
REFERENCE 1 (bases 1 to 880)
AUTHORS Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Wei,F., Begum,D.,
Frisch,D., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi
D.W., Fenton,R.D., Oates,R. and Main,D.
TITLE Development of a genetically and physically anchored EST resource
for barley genomics: Blumeria infected incompatible (M1a13)
seedling leaf cDNA library
JOURNAL Unpublished (2001)
COMMENT On Nov 17, 2000 this sequence version replaced gi:11195799.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 260
Seq primer: AATTAACCTCACTAAAGCG
High quality sequence stop: 798.
Location/Qualifiers
1..880
/organism="Hordeum vulgare"
/cultivar="C16155 (M1a13)"
/db_xref="taxon:4513"
/clone_lib="HV.CE0010619f"
/clone_lib="Hordeum vulgare seedling green leaf EST
library HVCNMA0004 (Blumeria challenged)"
/tissue_type="seedling green leaf"
/lab_host="TUC121"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: Xho1;
C.I. 16155 (M1a13) plants were greenhouse grown in the R
wise lab at Iowa State University, Ames, IA; 7 day old
green seedlings were challenged with isolate A27 (AVRMA13
) of Blumeria graminis f. sp. Hordei, and leaves were
harvested 20 and 24 hr post-inoculation and snap frozen;

uninoculated leaves were harvested 20 hr post-inoculation (Wei, Wise). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool. Equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture. One cDNA library was made, and 1 million pfu were in vivo excised to give plasmids SK(-) cDNA phagemids (Choi, Close). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Bequm, Palmett, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Math). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see

<http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinbols A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT 215 a 198 c 318 g 149 t

Query Match 3.0% Score 63.4; DB 10; Length 880;

Best Local Similarity 55.8%; Pred No. 4.1e-05;

Matches 121; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 605 gattactactaccacaaagagcgtccgcatctgttgatccatgacatgcatc 664

DB 74 GGCGTAGCGATACCGGACAGAGAGTCCCGGTACCTGTGTACACAGACCCAT 133

QY 665 gctacacacgcgcgaagcgtacatggtgtagctgtgtctacatgacagacc 724

DB 134 GGCGCTACCGCCGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 193

QY 725 gctgtagagatgctgacacacacacacacacacacacacacacacacacacac 784

DB 194 GGCCACGAGGAGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 253

QY 785 gactgcaagcagatatacaacgacgacacacacacacacacacacacacac 821

DB 254 CTTGACCGCGACTTCATCGCGAGCGCGCTCTTC 290

RESULT 6 465 bp mRNA linear EST 25-JAN-2001

LOCUS BG053739.1

DEFINITION RH122_9_B08.DL_A003 Rhizome2 (RH122) Sorghum propinquum cDNA, mRNA

ACCESSION BG053739

VERSION BG053739.1 GI:12509741

KEYWORDS EST

ORGANISM Sorghum propinquum

REFERENCE 1 (bases 1 to 465)

AUTHORS Cordomier-Pratt, M.-M., Gingle, A., Paterson, A., Sudman, M. and Pratt, L.H.

TITLE An EST database from Sorghum: Sorghum propinquum rhizomes

JOURNAL Unpublished (2000)

COMMENT Department of Botany

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 542 1805

Email: mmp@pratt.edu

Seq primer: JEN REV

High quality sequence stop: 421

POLYA-No.

FEATURES Location/Qualifiers

1..465

/organism="Sorghum propinquum"

/db_xref="taxon:132711"

/clone_lib="Rhizome2 (RH122)"

/note="Organ: Rhizomes; Vector: pBluescript II from lambda

zap II; Site:1: XhoI; Site:2: EcoRI; The library was made

from poly-A RNA in the cloning vector lambda zap II."

Clones to be sequenced were prepared by mass exsition."

BASE COUNT 72 a 128 c 193 g 72 t

ORIGIN

Query Match 2.6% Score 54.6; DB 10; Length 465;

Best Local Similarity 56.8%; Pred. No. 0.0046;

Matches 121; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

QY 609 tactactaccacaaagcagcgtgcccgcgtgtgtgtacatgacatgcatgctc 668

DB 452 TACCGGTACCCACACGTGACCGCGCGGACACTGTGTACACACACACGCGCTCGCC 393

QY 669 atcaccgcgcgaagcgtctacatggtgtagctgtgtctacatgacagaccgct 728

DB 392 CTCACCGCGCGCCACACTCTCGCGGCTCTCTGCGGCTTACGATGAGAAAGCGGAG 333

QY 729 gagatgacctgaacctcccccagcgtacgagatgtgatccctgtgttgcact 788

DB 332 GTCGACGCTCCATGATGATTTCTCTGCGACGAGCA---GACCTCCACCTCTGCTCGCG 276

QY 789 gccaaagcatatacaacgacgacacacacacacacacacacacacacacacacac 821

DB 275 GACCGACGCTTACGATGACGCGGCTGCTGTAC 243

RESULT 7

LOCUS AV426401

DEFINITION AV426401 Lotus japonicus young plants (two-week old) Lotus

ACCESSION AV426401

VERSION AV426401.1 GI:7785301

KEYWORDS EST

ORGANISM Lotus japonicus

REFERENCE 1 (bases 1 to 311)

AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.

TITLE Generation of 7137 non-redundant expressed sequence tags from a

legume, Lotus japonicus

JOURNAL DNA Res. 7 (2), 127-130 (2000)

COMMENT Contact: Yasukazu Nakamura

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yama 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: ynakamu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES

source

1..311

/organism="Lotus japonicus"

/db_xref="taxon:34305"

/clone_lib="MM066f04_r"

/dev_stage="young plants (two-week old)"

/note="Vector: pBluescript SK-; Site:1: EcoRI; Site:2: XhoI; isolate=Myakojima MG-20"

BASE COUNT 81 a 100 c 68 g 62 t

ORIGIN


```

/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/Note="Vector: pBluescript II SK+, Site 1: EcoRI, Site 2:
XhoI. The cDNA library was constructed from mRNA isolated
from 2-3 week old seedlings that were induced for symptoms
of SDS (Sudden Death Syndrome) disease by the
translocation of culture filtrate of Fusarium solani f.
sp. glycines (Plant Cell Report 18:375-380). Cultivar P.
567374 is partially resistant to the disease SDS. Plant
tissue (expanded leaves, folded leaves, and new shoots)
were collected at 1, 6, 24, and 48 hrs. after inoculation
and their mRNA pooled equally for cDNA construction. The
library was prepared using the Stratagene pBluescript II
SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with an XhoI restriction site. EcoRI
adaptors were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA insert is protected
from XhoI digestion via methylation during first strand
synthesis. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
E. coli Electromax DH10B host cells. Plants were inoculated
by Shuxian Li (Glen Hartman lab, University of Illinois).
Library was constructed by Steve Clough (Illa Vodkin lab,
University of Illinois)."
BASE COUNT      155 a      177 c      136 g      124 t
ORIGIN

```

```

Query Match      2.3%; Score 48.2; DB 10; Length 592;
Best Local Similarity 51.1%; Pred. No. 0.22;
Matches 113; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```

```

QY 601 acaagagattactactaccacacagcagcgtcccgatgcttgatcacatgacatg 660
      ||| || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 48 AGAGACATATCTCTACCCCAATATACACACAGGACCTATGTCATCATGACCATG 107
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 661 ccatgtccatcacccgcgagacgtacatggtcagctgggtgtctacatgacacag 720
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 108 CCATGGGGTTGACCGAGTCACCTTCTACCTGCGCTACGACCTCATCATGATTCGCC 167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 721 accgcgctgagatgcccctgaacctccacgagcgtacgagcgttgatatacccttg 780
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 168 ACCCTGAGATCGAGGACCGCTAGGCTTACCCAGCGGTGACGATTCGATGACGCTGA 227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 781 ttctgactgcacagatcacagcagcagcactctcttc 821
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 228 TCCTGTTCCATCCGACCTTCGACGACGCGTTGATCTTC 268
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
BG526127      766 bp      mRNA      linear      EST 16-NOV-2001
LOCUS        BG526127
DEFINITION   56-88 Stevia field grown leaf cDNA Stevia rebaudiana cDNA 5', mRNA
sequence.
ACCESSION   BG526127
VERSION     BG526127.1 GI:16949596
KEYWORDS    EST.
SOURCE      Stevia rebaudiana.
ORGANISM    Stevia rebaudiana.

```

```

REFERENCE    1 (bases 1 to 766)
AUTHORS      Brandie,J.E., Richman,A., Swanson,A.K. and Chapman,B.P.
TITLE        Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in
              diterpene synthesis
JOURNAL      Unpublished (2001)
COMMENT      Contact: Jim Brandie
              Genomics and Biotechnology
              Agriculture and Agri-Food Canada - SCFRC

```

```

1391 Sandford St., London, Ontario, CANADA, N5V 4T3
Tel: 519 457 1470
Fax: 519 457 3997
Email: brandiejeem@agr.ca
Seq primer: T3 promoter primer.
Location/Qualifiers
1. /766
/organism="Stevia rebaudiana"
/strain="751/1501"
/cultivar="landrace"
/db_xref="taxon:55670"
/clone_lib="Stevia field grown leaf cDNA"
/tissue_type="leaf"
/dev_stage="field grown, mid-size"
/lab_host="E. coli strain XL0LR"
/Note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI. This
cDNA library was constructed from polyA+ enriched mRNA
from field grown leaves. Mid-size actively growing leaves
were collected and pooled from several plants and frozen
immediately after harvesting in liquid nitrogen. The cDNA
was prepared using an XhoI-poly(dT) linker-primer. An
EcoRI adapter was ligated to the blunt end cDNA and the
products were digested with EcoRI and XhoI enabling
directional cloning into the lambda ZAP Express vector.
The library was amplified using the host strain XL1-Blue
MRP. Mass excision of the library was performed to
obtain pBK-CMV phagemid clones in the host strain XL0LR.
Single pass DNA sequencing was performed using the T3
promoter primer: 5' ATTACCTCCCTCAAGGGA 3'. This library
was constructed by Alex Richman."
BASE COUNT      203 a      192 c      174 g      191 t
ORIGIN

```

```

Query Match      2.3%; Score 47.8; DB 10; Length 766;
Best Local Similarity 52.9%; Pred. No. 0.33;
Matches 100; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```

```

QY 601 acaagattactactaccacacagcagcgtcccgatgcttgatcacatgacatg 660
      ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 64 AAANGMCTATGACCTTACACACACACACAGCCGGAACCTTTATTCATGATGACG 123
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 661 ccatgtccatcacccgcgagacgtacatggtcagctgggtgtctacatgacacag 720
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 124 CCATGGGGTTGACCGAGTCACCTTCTACCTGCGCTACGACCTCATCATGATTCGCC 183
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 721 accgcgctgagatgcccctgaacctccacgagcgtacgagcgttgatatacccttg 780
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 184 ATCAGATGTCGAGACCCCACTTGCGCTCTTGCTGTGATGATGCGCGCTTG 243
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 781 ttctgactg 789
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 244 TCCTATTTG 252
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 11
A2714471      286 bp      DNA      linear      GSS 24-JAN-2001
LOCUS        A2714471
DEFINITION   RPCI-24-138E2.TJ RPCI-24 Mus musculus genomic clone RPCI-24-138E2,
DNA sequence.
ACCESSION   A2714471
VERSION     A2714471.1 GI:12450418
KEYWORDS    GSS.
SOURCE      Mus musculus.
ORGANISM    Mus musculus.

```

```

REFERENCE    1 (bases 1 to 286)
AUTHORS      Mammalla; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
              Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akhmet,B., Levins,M.,
              Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregorgis,E.,
              Russell,D., de Jong,P. and Fraser,C.M.
TITLE        Mouse BAC End Sequences from Library RPCI-24
JOURNAL      Unpublished (1999)

```


COMMENT

Other_GSSs: RPCI-24-138E2.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdjong@tigr.org). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/ordering.htm>). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 138 row: E column: 2
 Seq primer: SP6
 Class: BAC ends.

FEATURES

source Location/Qualifiers
 1..286

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-24-138E2"
 /clone_1ib="RPCI-24"
 /sex="Male"
 /cell_type="Spleen/Brain"
 /note="Vector: pTARBAC1; Site_1: BamHI, Site_2: BamHI, RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 71 a 79 c 67 g 69 t

ORIGIN

Query Match 2.2%; Score 46.2; DB 12; Length 286;
 Best Local Similarity 52.3%; Pred. No. 0.46;

Matches 102; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

OY 543 tcttctcgcagctcccttgatggttggtgcgagacactcccgctgagtagtac 602

Db 36 tcttctccttactgacgagatgccaagatgaagacatctctggttctccccaat 95

OY 603 aaggaattactactaccacacagcagcagctgcgcgcatgcttgtaaccatgac 662

Db 96 GGGAGAGAGCTGAGCCCAACAGCAGCGGATCTCAGTGTGGATGATGATCTTTC 155

OY 663 atgtccatcacgcgcgagaaacctacatggttcagagctggtgtctacatgaccagac 722

Db 156 tttacccttcacatctacacacgacacatcgcgcatatctacagtggtgtc 215

OY 723 ccgagctgagatgac 737

Db 216 ACCGCTGAGCAGCGC 230

RESULT 12

LOCUS

BF258534 599 bp mRNA linear EST 23-OCT-2001

DEFINITION HVSMBF0015P21f Hordeum vulgare seedling root EST library HVCDA0007

(Etiolated and unstressed) Hordeum vulgare cDNA clone

Accession HVSMBF0015P21f, mRNA sequence.

VERSION

KEYWORDS

SOURCE

ORGANISM

barley.
 Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 1 (bases 1 to 599)

REFERENCE

Wing, R., Close, T.J., Kleinbols, A., Wise, R., Begum, D., Frisch, D., Yu,
 Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton,
 R.D., Oates, R. and Main, D.

TITLE

JOURNAL

COMMENT

Development of a genetically and physically anchored EST resource
 for barley genomics: Morex unstressed seedling root cDNA library
 Unpublished (2001)
 On Nov 16, 2000 this sequence version replaced gi:11187647.
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hg bases = 451
 Seq primer: AATTAACCTCACTAAGG
 High quality sequence stop: 552.

FEATURES

source

Location/Qualifiers
 1..599
 /organism="Hordeum vulgare"
 /cultivar="Morex"
 /db_xref="taxon:4513"
 /clone="HVSMBF0015P21f"
 /clone_1ib="Hordeum vulgare seedling root EST library
 HVCDA0007 (Etiolated and unstressed)"
 /tissue_type="Seedling root"
 /lab_host="TTC121"

/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
 Seeds were surface sterilized then germinated under axenic
 conditions in the dark at room temperature on filter paper
 with water, nystatin and cefotaxime in covered
 crystallization dishes. Five-day old seedling roots were
 then harvested, total RNA was prepared, poly(A) RNA was
 purified, one primary unamplified cDNA library was made,
 and 1 million pfu were in vivo excised to give plasmid
 SK(-) cDNA phagemids. These steps were performed in the TU
 close laboratory at the University of California,
 Riverside (Choi, Close, Fenton). Phagemids were plated and
 picked at the Clemson University Genomics Institute (CUGI)
 (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
 preparations, DNA sequencing and sequence analysis were
 performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
 , Rambo, Main). The sequence has been trimmed to remove
 vector sequence and contains a minimum of 100 bases of
 phred value 20 or above. For more details on library
 preparation and sequence analysis see
<http://www.genome.clemson.edu/projects/barley>. To order
 this clone see <http://www.genome.clemson.edu/orders> Also
 see Close T.J., Wing R., Kleinbols A., Wise R. (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (<http://wheat.pw.usda.gov/gpages/bgn/31/cover.html>).

BASE COUNT 96 a 215 c 206 g 82 t

ORIGIN

Query Match 2.2%; Score 46.2; DB 10; Length 599;
 Best Local Similarity 56.8%; Pred. No. 0.73;

Matches 105; Conservative 0; Mismatches 78; Indels 2; Gaps 1;

OY 953 aacgctcagccttgctatgctcaacgctgacgcgcgaagatgacgcttcctc 1012

Db 77 AACGGCAAGGCTTGCGCTTCGCCGTGCGCGCCGCCCTACCGCTCCCACTCC 136

OY 1013 aacgctcagccttcagccttcgctgctatctgctactc-ctgagatcagaga 1070

Db 137 AACGCCAGCAGCGCGCCTTCCTCCGCTTCCTCCGCGCGCTTCGCTTGTGAC 196

OY 1071 ccagactcccttcagcagatcgcgcgctgacggtgctgcttgagagccctgtgaca 1130

Db 197 GTCGCTCCGACACGCTGCTACCTGCGCGCGCGCGCCGACGAGGAGTTCTGCTGCG 256

OY 1131 ctgac 1135

Db 257 CCGCTC 261

Mon Oct 7 11:28:44 2002

us-09-656-640a-3.rst

Page 10

Db 483 CACGGCTGAGAGGC 498

Search completed: October 5, 2002, 18:54:25
Job time: 3515 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 20:20:09 ; Search time 92.23 Seconds
(without alignments)
5579.553 Million cell updates/sec

Title: US-09-656-640A-3
Perfect score: 2095
Sequence: 1 cagctcggtactactctc.....gaataagtgtgtgtctaa 2095

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2-6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2-6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2-6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2-6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2-6/ptodata/1/ina/PCUTUS_COMB.seq:*
6: /cgn2-6/ptodata/1/ina/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2095	100.0	2095	US-09-401-476-3	Sequence 3, Appli
2	1946.2	92.9	1958	US-09-401-476-1	Sequence 1, Appli
3	50	2.4	7218	US-08-232-463-14	Sequence 14, Appli
4	39.4	1.9	2071	US-07-923-724-1	Sequence 1, Appli
5	39.4	1.9	2071	US-08-609-426A-1	Sequence 1, Appli
6	39.4	1.9	2071	US-08-374-652C-3	Sequence 3, Appli
7	37.6	1.8	289	US-09-007-005-17	Sequence 17, Appli
8	37.6	1.8	289	US-09-244-796-17	Sequence 17, Appli
9	35.2	1.7	1704	US-08-528-199-2	Sequence 2, Appli
10	35.2	1.7	1704	US-08-528-199-5	Sequence 5, Appli
11	34.8	1.7	1637	US-08-966-316-10	Sequence 10, Appli
12	34.8	1.7	2139	US-08-286-870A-7	Sequence 7, Appli
13	34.2	1.6	783	US-08-446-922-5	Sequence 5, Appli
14	34.2	1.6	783	US-08-249-189-1	Sequence 1, Appli
15	34.2	1.6	783	US-08-484-624A-1	Sequence 1, Appli
16	34.2	1.6	783	US-08-477-733B-1	Sequence 1, Appli
17	34.2	1.6	783	US-09-088-913A-1	Sequence 1, Appli
18	34.2	1.6	783	US-08-769-819-1	Sequence 1, Appli
19	34.2	1.6	783	US-08-770-974-1	Sequence 1, Appli
20	34.2	1.6	878	PCT-US93-10034-5	Sequence 5, Appli
21	34.2	1.6	878	US-08-249-189-22	Sequence 22, Appli
22	34.2	1.6	878	US-08-484-624A-22	Sequence 22, Appli
23	34.2	1.6	878	US-08-477-733B-22	Sequence 22, Appli
24	34.2	1.6	878	US-09-088-913A-22	Sequence 22, Appli
25	34.2	1.6	878	US-08-769-819-22	Sequence 22, Appli
26	34.2	1.6	878	US-08-770-974-22	Sequence 22, Appli
27	34.2	1.6	5392	US-08-403-852D-1	Sequence 1, Appli

c	28	34.2	1.6	5392	3	US-08-510-646B-1	Sequence 1, Appli
c	29	34.2	1.6	5392	4	US-09-231-818-1	Sequence 1, Appli
c	30	34	1.6	362	1	US-08-486-895-2	Sequence 1, Appli
c	31	34	1.6	443	1	US-08-486-895-1	Sequence 1, Appli
c	32	34	1.6	1248	4	US-09-105-537-7	Sequence 2, Appli
c	33	34	1.6	5970	3	US-09-320-878-21	Sequence 21, Appli
c	34	34	1.6	13613	4	US-09-105-537-3	Sequence 3, Appli
c	35	34	1.6	30001	1	US-08-125-468-1	Sequence 1, Appli
c	36	34	1.6	30001	2	US-08-474-933-1	Sequence 1, Appli
c	37	34	1.6	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	38	33.8	1.6	3417	2	US-08-464-402-1	Sequence 1, Appli
c	39	33.8	1.6	3417	4	US-09-054-775C-1	Sequence 1, Appli
c	40	33.2	1.6	752	4	US-08-818-112-2	Sequence 2, Appli
c	41	33.2	1.6	752	4	US-08-818-111-2	Sequence 2, Appli
c	42	33.2	1.6	752	4	US-09-056-556-2	Sequence 2, Appli
c	43	33.2	1.6	1086	1	US-08-255-670A-3	Sequence 3, Appli
c	44	33.2	1.6	1624	1	US-08-255-670A-1	Sequence 1, Appli
c	45	33	1.6	1908	1	US-08-173-508-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-401-476-3
Sequence 3, Application US/09401476
Patent No. 6168936
GENERAL INFORMATION:
APPLICANT: Wang, Huang
TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes
FILE REFERENCE: GC584
CURRENT APPLICATION NUMBER: US/09/401,476
CURRENT FILING DATE: 1999-09-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 3
LENGTH: 2095
TYPE: DNA
ORGANISM: Stachybotrys chararum
US-09-401-476-3

Query Match 100.0%; Score 2095; DB 4; Length 2095;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2095; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	cagctcggtactactctcgtctctcttgacaataatctacacatcgctctca	60
DB	1	cagctcggtactactctcgtctctcttgacaataatctacacatcgctctca	60
QY	61	attcaaacgatacaatgatacgaacgatacgagcggtggtcttggtctgt	120
DB	61	attcaaacgatacaatgatacgaacgatacgagcggtggtcttggtctgt	120
QY	121	gattcgagcgagctctgctgatacgacagatcggtctgctgatacgacagatc	180
DB	121	gattcgagcgagctctgctgatacgacagatcggtctgctgatacgacagatc	180
QY	181	cggttcacacagagcgagcgacgactgactctccctcggtctgtaagagtcct	240
DB	181	cggttcacacagagcgagcgacgactgactctccctcggtctgtaagagtcct	240
QY	241	gcgcattccctctctgaagcgcccaagtagtaagtaattctataggctagcagc	300
DB	241	gcgcattccctctctgaagcgcccaagtagtaagtaattctataggctagcagc	300
QY	301	aggttgtaattatgagctgacggtcccaacccaacgaactgagaggaactgttact	360
DB	301	aggttgtaattatgagctgacggtcccaacccaacgaactgagaggaactgttact	360
QY	361	acgagatgagatgagctgacgtctctccacacagatctacatctgagcgagcaca	420
DB	361	acgagatgagatgagctgacgtctctccacacagatctacatctgagcgagcaca	420

```
OY 421 tgggttgatagcgtgcatgctccagagacctaaccatcgtcttcctcgttggcaattgaga 480
|||||
Db 421 tgggttgatagcgtgcatgctccagagacctaaccatcgtcttcctcgttggcaattgaga 480
OY 481 gttgttcgcgtctcgttgagacagcgagagaaacacactctcccaacagctccacttgagc 540
|||||
Db 481 gttgttcgcgtctcgttgagacagcgagagaaacacactctcccaacagctccacttgagc 540
OY 541 gctctctctcgaagctccctcttgatgtgtggcctgagagacctaaccagcttgagagat 600
|||||
Db 541 gctctctctcgaagctccctcttgatgtgtggcctgagagacctaaccagcttgagagat 600
OY 601 acaagattactactaaccacacagcagcgctgcgcgcagtcttgatgatacagacatg 660
|||||
Db 601 acaagattactactaaccacacagcagcgctgcgcgcagtcttgatgatacagacatg 660
OY 661 ccatgttccatccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 720
|||||
Db 661 ccatgttccatccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 720
OY 721 acccgcttgagagatgcccctggaacccccccagcgctacgagcttgatgatacccttgg 780
|||||
Db 721 acccgcttgagagatgcccctggaacccccccagcgctacgagcttgatgatacccttgg 780
OY 781 tcttgactgccaagcgatacaacgcagacgcaactctcttccacacaaatgagagatt 840
|||||
Db 781 tcttgactgccaagcgatacaacgcagacgcaactctcttccacacaaatgagagatt 840
OY 841 ccaagttctgggggtgagctgtatcttaagtgtaagtgagccattgagatgcttcaagtc 900
|||||
Db 841 ccaagttctgggggtgagctgtatcttaagtgtaagtgagccattgagatgcttcaagtc 900
OY 901 ctagaagatcagatglatgaaattgtgcatgctcctaaccagtgctataccaacagtgta 960
|||||
Db 901 ctagaagatcagatglatgaaattgtgcatgctcctaaccagtgctataccaacagtgta 960
OY 961 gctctgagcctatgctcaacgctgacgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1020
|||||
Db 961 gctctgagcctatgctcaacgctgacgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1020
OY 1021 cgtctcaacgctcttgcgtctgtatcttgatcctctgagagattagagacagacattcc 1080
|||||
Db 1021 cgtctcaacgctcttgcgtctgtatcttgatcctctgagagattagagacagacattcc 1080
OY 1081 ctcccaaggtcattgacgcgcgtgacgctgctgtcttgagagccctcttgacactgaactct 1140
|||||
Db 1081 ctcccaaggtcattgacgcgcgtgacgctgctgtcttgagagccctcttgacactgaactct 1140
OY 1141 gtacactctatgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1200
|||||
Db 1141 gtacactctatgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1200
OY 1201 gtccatgatatccgacacactcctcggtgctgacgagtcctcgtgttgagccctgagattga 1260
|||||
Db 1201 gtccatgatatccgacacactcctcggtgctgacgagtcctcgtgttgagccctgagattga 1260
OY 1261 taacactgacaaggtcagcagatcgtctgctgataagtcctctgagtcgcgcgcgcgcgcgcgc 1320
|||||
Db 1261 taacactgacaaggtcagcagatcgtctgctgataagtcctctgagtcgcgcgcgcgcgcgcgc 1320
OY 1321 tgaagtgctgcgaacccctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1380
|||||
Db 1321 tgaagtgctgcgaacccctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1380
OY 1381 aaaccccaactgatacagcagacttcaactctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1440
|||||
Db 1381 aaaccccaactgatacagcagacttcaactctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1440
OY 1441 cggagattacattcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1500
|||||
Db 1441 cggagattacattcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1500
```

```
OY 1501 tgaatctggcagacttgagaaacactccaaacggttgagactcaaccctgttacaattacact 1560
|||||
Db 1501 tgaatctggcagacttgagaaacactccaaacggttgagactcaaccctgttacaattacact 1560
OY 1561 cgttgacttcgcagcctcttctcgttccactgcccgttgagctcagagccttaagaagctgc 1620
|||||
Db 1561 cgttgacttcgcagcctcttctcgttccactgcccgttgagctcagagccttaagaagctgc 1620
OY 1621 tggctcacaagatgctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1680
|||||
Db 1621 tggctcacaagatgctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1680
OY 1681 cgtctccttcccgtaagctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1740
|||||
Db 1681 cgtctccttcccgtaagctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1740
OY 1741 acaagtggtctaatatgctgactgcgcacacacccgataccagagacacagacatgctg 1800
|||||
Db 1741 acaagtggtctaatatgctgactgcgcacacacccgataccagagacacagacatgctg 1800
OY 1801 gctgcttcaatgctacgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1860
|||||
Db 1801 gctgcttcaatgctacgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1860
OY 1861 cccatgagagcctctctgagagcccgccctctcctcctcctcctcctcctcctcctcctcctc 1920
|||||
Db 1861 cccatgagagcctctctgagagcccgccctctcctcctcctcctcctcctcctcctcctcctc 1920
OY 1921 ggtgacttcaagagccttgacacactgacacacacacacacacacacacacacacacacacac 1980
|||||
Db 1921 ggtgacttcaagagccttgacacactgacacacacacacacacacacacacacacacacacac 1980
OY 1981 taagcccaagcctgagatgctgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2040
|||||
Db 1981 taagcccaagcctgagatgctgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2040
OY 2041 tggacagcagatatacgtattttaagaaactcttgaaatgctgtgtgtctaa 2095
|||||
Db 2041 tggacagcagatatacgtattttaagaaactcttgaaatgctgtgtgtctaa 2095

RESULT 2
US-09-401-476-1
; Sequence 1, Application US/09401476
; Patent No. 6168936
; GENERAL INFORMATION:
; APPLICANT: Wang, Huming
; TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes
; FILE REFERENCE: GC584
; CURRENT APPLICATION NUMBER: US/09/401,476
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1958
; TYPE: DNA
; ORGANISM: Stachybotrys chararum
US-09-401-476-1

Query Match 92.9%, Score 1946.2, DB 4, Length 1958:
Best Local Similarity 99.8%, Pred. No. 0,
Matches 1948, Conservative 0, Mismatches 3, Indels 0, Gaps 0:

OY 67 aacgatacaatgatacagcacaagcctacatcgagacgctgctcgggacctgtcgtatcgg 126
|||||
Db 3 atccatcaacatgatacagcacaagcctacatcgagacgctgctcgggacctgtcgtatcgg 62
OY 127 cggagacctgtcgcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 186
|||||
Db 63 cggagacctgtcgcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 122
OY 187 caccagaagcagacgacgctgagctcctccctggccttgtaagagtgctccttgccgct 246
|||||
```

[illegible]

Dh	1203	tgaaagatgacatgcatctgctgctgtgaatgaagtcctcttgatgctgcgcgcgaactcttgagt	1262
QY	1327	gctcgcaacccctccgagatgctcctcttcccgaagcgcgcaactctgagaccccgcaaaccc	1386
Dh	1263	gctcgcaacacccctccgagatgctcctcttcccgaagcgcgcaactctgagaccccgcaaaccc	13322
QY	1387	cactcgatgacagaaacttccactctgcccgtctctaatagacagctggaacaatccacgagat	1446
Dh	1323	cactcgatgacagaaacttccactctgcccgtctctaatagacagctggaacaatccacgagat	1382
QY	1447	tacctctcgagatgctcgaagaacgctctgtctccgcaatgctgcgcgcgcaactgttgaat	1506
Dh	1383	tacctctcgagatgctcgaagaacgctctgtctccgcaatgctgcgcgcgcaactgttgaat	1442
QY	1507	ctgcgagacttgagaaacaactccacagctctgagatccaaacctgtctacatccactctgtga	1566
Dh	1443	ctgcgagacttgagaaacaactccacagctctgagatccaaacctgtctacatccactctgtga	1502
QY	1567	cttccgagctccttctcgtctccactcgaatcgccgtgagatcgagccttatgagctgtgtct	1626
Dh	1503	cttccgagctccttctcgtctccactcgaatcgccgtgagatcgagccttatgagctgtgtct	1562
QY	1627	caagagatgtctctgctcgtccgtccgtctcgatgagttctcatgtcttgagcgccactaaagctcc	1686
Dh	1563	caagagatgtctctgctcgtccgtccgtctcgatgagttctcatgtcttgagcgccactaaagctcc	1622
QY	1687	tttcccgtaaagtctctcgacttctacccaagaatgcttttccactatgcttaacatcttaagaat	1746
Dh	1623	tttcccgtaaagtctctcgacttctacccaagaatgcttttccactatgcttaacatcttaagaat	1682
QY	1747	ggtgtctcaatgtctgcaactctgccaacactgataccacgagagacacagacaatgactgtct	1806
Dh	1683	ggtgtctcaatgtctgcaactctgccaacactgataccacgagagacacagacaatgactgtct	1742
QY	1807	ttcaagtctcaactttctccgctgagctatgctctacaactaaacagagtttatgtgacccatg	1866
Dh	1743	ttcaagtctcaactttctccgctgagctatgctctacaactaaacagagtttatgtgacccatg	1802
QY	1867	gagagctctcgaagcgcccccctctccactccctctcgaagaatctcgaagaatgctcgggttac	1926
Dh	1803	gagagctctcgaagcgcccccctctccactccctctcgaagaatctcgaagaatgctcgggttac	1862
QY	1927	ttcagcgagctctgcatcaactgacgcgcatctcagagagatgctagctttcaacccctaagc	1986
Dh	1863	ttcagcgagctctgcatcaactgacgcgcatctcagagagatgctagctttcaacccctaagc	1922
QY	1987	caagctctatgataatgctccgcttgagagataa	2017
Dh	1923	caagctctatgataatgctccgcttgagagataa	1953

RESULT 3
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT29pt-F15
US-08-232-463-14

```

```

Query Match          2.4% Score 50; DB 1; Length 7218;
Best Local Similarity 0.5%; Pred. No. 0.00012;
Matches 2; Conservative 221; Mismatches 141; Indels 0; Gaps 0;

```

```

OY 1539 ctccacctgtacatctcacctcgtgacctccgagctcttcgtctccactgcgcgtg 1598
DB 1074 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1133
OY 1599 gagtcgagcctatagagctgctgctcgaagatgctgcgcgcgcgcgcgcgcgcgc 1658
DB 1134 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1193
OY 1659 ttgtctatgttagagccactacgctcctccgtaagttcgccttaccactgactg 1718
DB 1194 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1253
OY 1719 gtttcaactcactgatacaagtgctgctacatgctgacatgcacaaactgact 1778
DB 1254 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1313
OY 1779 ccacgagaccacgacatgctgcttcaatgctcactgcttcgctgactatgacta 1838
DB 1314 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1373
OY 1839 caactacacgagctcatgaccacatgagcctcctctgagagcccgcccttccctc 1898
DB 1374 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1433
OY 1899 cggga 1902
DB 1434 ygtta 1437

```

```

RESULT 4
US-07-923-724-1
Sequence 1, Application US/07923724
Patent No. 5780292
GENERAL INFORMATION:
APPLICANT: Nevalainen, Helena K.M.
APPLICANT: Paloneimo, Marja T.
APPLICANT: Miettinen-Oinonen, Arja S.K.
APPLICANT: Torkkeli, Tuula K.
APPLICANT: Cantrell, Michael
APPLICANT: Piddington, Christopher S.
APPLICANT: Ramdasek, John A.

```

```

APPLICANT: Turunen, Marja K.
APPLICANT: Fagerstr m, Richard B.
TITLE OF INVENTION: Production of Phytase Degrading Enzymes
TITLE OF INVENTION: in Trichoderma
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSER: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/923,724
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/496,155
FILING DATE: 19-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/044,077
FILING DATE: 29-APR-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 8610600
FILING DATE: 30-APR-1986
ATTORNEY/AGENT INFORMATION:
NAME: Cimbal, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 1050,0240004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2071 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: join(136..915, 970..1089, 1142..1245, 1305..1737)
US-07-923-724-1

```

```

Query Match          1.9% Score 39.4; DB 1; Length 2071;
Best Local Similarity 45.6%; Pred. No. 0.088;
Matches 139; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

```

```

OY 455 catctcgtctcgtgagactgagatgtgtgcgcttcgagctcggaacgagagagac 514
DB 444 CAGCATCAACTACTGATGATACAGAGGACCTGGCCTTCTGAAGACTGACCTACTA 503
OY 515 ctctccaaacagctccactgacagcgtctctctcgtcagctcccttgatgtgtg 574
DB 504 CGTCCCTATGATGCTACTACTACAGCGCGAGACGAGCGGCGCCCTACGCGGTTTGT 563
OY 575 tgaagacactaccagcctgagctgacaaagattactactaccccaacagcagctgc 634
DB 564 GAGCCGTTACACACATGGCAACGATTACAGAGCTGCTACGCGCACCTCTGCAACGCTGA 623
OY 635 ccgcatgcttctgtaaccatgacatgcatgtccatcaccgagcaggaagcctaatagg 694
DB 624 GACGCTGCGCCCTCTTTCTAGTGCTACGAGCTGATCGAGAGCGCCGCAAGTT 683
OY 695 tcaagctgtgtctacatgatccagagaccgctgagatgctccgaacctcccaagcgg 754
DB 684 CGGTAGAGGTTCTTTGCTGCTACACTACTTCCACCAACGCTGCGCTCAACATATCTCCGA 743

```


Oy 836 gg 837
Db 1350 GG 1351

RESULT 13

US-08-446-922-5/C
; Sequence 5, Application US/08446922
; Patent No. 5716805
; GENERAL INFORMATION:
; APPLICANT: Spriggs, Melanie
; APPLICANT: Srinivasan, Subhashini
; TITLE OF INVENTION: Methods of Preparing Soluble, Oligomeric
; NUMBER OF INVENTION: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Operating System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,922
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/107,353
; FILING DATE: 08-13-93
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 1003-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Mouse
; STRAIN: CD40-L
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..780
; US-08-446-922-5

Query Match 1.6%; Score 34.2; DB 1; Length 783;
Best Local Similarity 43.6%; Pred.No.1.8;
Matches 153; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

Oy 1358 gagggcggaactggagcccggaacccgaactgtagcagagacttcaacttggcgagt 1417
Db 774 GAGTAAAGCCAAAGATGAGAGACCACTCTGTGATCACTTGTGCTTCAATCAGTCA 715
Oy 1418 gctaatgacagtgtagcagatcaacggagttacctctctgtagtgtagaagcgttgc 1477
Db 714 GACAAACACAGAGAGACCAAGCTTTGATTAATCAACACTCCGCCCAAGTACAGACTG 655
Oy 1478 cgcaatgtgccccgagacactgttagatcttggcgacttggagaaacactccacaggttgg 1537

Db 654 CTCGAAAGCTGGGAGGAGACTGTGGTATTTCGCCCTTGAGTAGATTCCTTCAGATCC 595
Oy 1538 actcaaccctgttcacattcaactctgttgactctcgagtccttccttcgctcaactcccg 1597
Db 594 ACTGCTGGGCTTCACCGACAGAGCCGACATGATGATGGCTTACACTGGAAGGCTCCGAT 535
Oy 1598 gagtcgagccttatgagagctgtctgtctcaagatgtgtctgctgctgctgctgag 1657
Db 534 AGAGCAGAGAGTACTGTGATGTAGACATAATAGAGTCCCTCTTTTACCGTCAAGCTG 475
Oy 1658 gtgtctatgttgagggccactacgctccttcctccgtaagttcgcctt 1708
Db 474 TTCCCATTTTCAAGCATTAACCAAGTTGCTTTTCATGTATATATCTTT 424

RESULT 14

US-08-249-189-1/C
; Sequence 1, Application US/08249189
; Patent No. 5961974
; GENERAL INFORMATION:
; APPLICANT: ARMITAGE, RICHARD
; APPLICANT: FANSLAW, WILLIAM
; APPLICANT: SPRIGGS, MELANIE
; APPLICANT: SRINIVASAN, SUBHASHINI
; APPLICANT: GIBSON, MARYLOU
; TITLE OF INVENTION: NOVEL CYTOKINE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: IMMUNEX CORPORATION
; STREET: 51 UNIVERSITY STREET
; CITY: SEATTLE
; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Operating System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/249,189
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/969,703
; FILING DATE: October 23, 1992
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/805,723
; FILING DATE: December 5, 1991
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/783,707
; FILING DATE: October 25, 1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2802-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 2065870430
; TELEFAX: 2065870606
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:

ORGANISM: MOUSE
IMMEDIATE SOURCE:
CLONE: CD40-L
FEATURE:
NAME/KEY: CDS
LOCATION: 1..783
US-08-249-189-1

Query Match 1.6% Score 34.2; DB 2; Length 783;
Best Local Similarity 43.6%; Pred. No. 1.8;
Matches 153; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 1358 gagggcggaactggagaccgcaaccccaactgacagacattcaactcggccgt 1417
DB 774 GAGTAAGCAAAAGATGAGAGCCAACTGTGATCACTTGCTGGCTGACGATT 715
QY 1418 gctaagacagtgagacaatcaacggagttactctcgagatgacagacgtctgc 1477
DB 714 GACAAACACAGAGACACAGCTTGTAAATTCAAACACTCCGCCAAGTGACAGACTGCG 655
QY 1478 cgcaatgtgccccgagacactgtgagatctgagacattgagacaactccaaggttgg 1537
DB 654 CTCGAAAGCTGGGAGAGACTGTGGTATTTCCTTGAGTAAGATTCTCTCAGATCC 595
QY 1538 actcaactgttcaactcaactcgttgaactcagagtccttccgtccactgcccgt 1597
DB 594 ACTGCTGGCTTCACGACAGAGCCGACATGATGAGCGCTTGACTGGAAGGCTCCGATT 535
QY 1598 ggaatcgagcctatgagagcgtgctgctcaagagatgtgtgtgctgagctgag 1657
DB 534 AGAGCAGAGGAGCTTGTAGTACATAATAGAGTCTTCTTTAAACCGTACGCTG 475
QY 1658 gtgctatgtgagagcccaactcgtccttccgtaagtctgcctt 1708
DB 474 TTTCATTTTCAACGATTACCAAGTGTCTTTTCATGTGATATAATCTTT 424

RESULT 15
US-08-484-624A-1/c
Sequence 1, Application US/08484624A
Patent No. 5962406
GENERAL INFORMATION:
APPLICANT: ARMITAGE, RICHARD
APPLICANT: FANSLAW, WILLIAM
APPLICANT: SPRIGGS, MELANIE
APPLICANT: SRINIVASAN, SUBHASHINI
APPLICANT: GIBSON, MARYLOU
APPLICANT: MORRIS, ARVIA E.
APPLICANT: MCGREW, JEFFERY
TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,624A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,733
FILING DATE: June 07, 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/969,703
FILING DATE: October 23, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/805,723
FILING DATE: December 5, 1991
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/783,707
FILING DATE: October 25, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2802-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2065870430
TELEFAX: 2065870606
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 783 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: MOUSE
IMMEDIATE SOURCE:
CLONE: CD40-L
FEATURE:
NAME/KEY: CDS
LOCATION: 1..783
US-08-484-624A-1

Query Match 1.6% Score 34.2; DB 2; Length 783;
Best Local Similarity 43.6%; Pred. No. 1.8;
Matches 153; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 1358 gagggcggaactggagaccgcaaccccaactgagacagacattcaactcggccgt 1417
DB 774 GAGTAAGCAAAAGATGAGAGCCAACTGTGATCACTTGCTGGCTGACGATT 715
QY 1418 gctaagacagtgagacaatcaacggagttactctcgagatgacagacgtctgc 1477
DB 714 GACAAACACAGAGACACAGCTTGTAAATTCAAACACTCCGCCAAGTGACAGACTGCG 655
QY 1478 cgcaatgtgccccgagacactgtgagatctgagacattgagacaactccaaggttgg 1537
DB 654 CTCGAAAGCTGGGAGAGACTGTGGTATTTCCTTGAGTAAGATTCTCTCAGATCC 595
QY 1538 actcaactgttcaactcaactcgttgaactcagagtccttccgtccactgcccgt 1597
DB 594 ACTGCTGGCTTCACGACAGAGCCGACATGATGAGCGCTTGACTGGAAGGCTCCGATT 535
QY 1598 ggaatcgagcctatgagagcgtgctgctcaagagatgtgtgtgctgagctgag 1657
DB 534 AGAGCAGAGGAGCTTGTAGTACATAATAGAGTCTTCTTTAAACCGTACGCTG 475
QY 1658 gtgctatgtgagagcccaactcgtccttccgtaagtctgcctt 1708
DB 474 TTTCATTTTCAACGATTACCAAGTGTCTTTTCATGTGATATAATCTTT 424

Search completed: October 5, 2002, 20:20:55
Job time: 8530 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:54:56 ; Search time 19.26 Seconds

(Without alignments)
1172.040 Million cell updates/sec

Title: US-09-656-640A-2

Perfect score: 3114

Sequence: 1 MISQAGNAGLAVIGGSS.....IQEMAFNPYAQADDAAE 583

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2010	64.5	572	1	BLRO_MYRVE
2	731.5	23.5	513	1	COTR_BACSU
3	536.5	17.2	642	1	PHSA_STRAT
4	412.5	13.2	516	1	YACK_ECOLI
5	319	10.2	470	1	SUFI_ECOLI
6	310	10.0	470	1	SUFI_SALTY
7	247.5	7.9	622	1	FETS_YEAST
8	247	7.9	311	1	SUFI_HAEPN
9	219.5	7.0	591	1	LACI_CRYPA
10	218.5	7.0	624	1	FETS_CANAL
11	199.5	6.4	527	1	LACS_TRAVE
12	198	6.4	531	1	LAC2_THACU
13	196.5	6.3	622	1	YAK8_SCHPO
14	195.5	6.3	520	1	LACI_TRAVI
15	193	6.2	529	1	LACI_PLEOS
16	192	6.2	619	1	LACI_NEUCR
17	190	6.1	636	1	FETS_YEAST
18	188.5	6.1	533	1	LAC2_PLEOS
19	188	6.0	548	1	LAC1_PHIRA
20	188	6.0	572	1	LAC3_THACU
21	187.5	6.0	527	1	LACS_TRAVI
22	187.5	6.0	621	1	LAC2_PODAN
23	186.5	6.0	520	1	LAC4_TRAVE
24	185.5	6.0	486	1	LAC1_BOTCI
25	185.5	5.9	520	1	LAC4_TRAVI
26	183	5.9	619	1	LAC2_NEUCR
27	182.5	5.9	587	1	ASO_CUCSA
28	178.5	5.7	520	1	LAC1_CORNI
29	178	5.7	609	1	COPA_PSESM
30	177	5.7	519	1	LAC2_TRAVI
31	176.5	5.7	576	1	LAC1_THACU
32	174	5.6	579	1	ASO_CUCMA
33	173.5	5.6	599	1	LAC2_THACU

34	172.5	5.5	605	1	PCOA_ECOLI	047452	escherichia
35	167	5.4	519	1	LAC2_TRAVE	012718	trametes ve
36	166	5.3	520	1	ASO_CUCPM	P37064	cucurbita p
37	165	5.3	520	1	LAC1_AGABI	012541	agaticus bl
38	160	5.1	520	1	LAC2_AGABI	012542	agaticus bl
39	160	5.1	608	1	YD56_YEAST	004399	saccharomyc
40	151.5	4.9	473	1	LAC1_TRAVI	099049	trametes vl
41	147.5	4.7	609	1	LAC1_EMEI	P17489	emeritella
42	139	4.5	578	1	ASO_TOBAC	040588	nicotiana t
43	123.5	4.0	841	1	MYFC_YEREN	P33408	yersinia en
44	118	3.8	374	1	NIR_RHOSH	053239	rhodobacter
45	116	3.7	454	1	CBPH_LOPAM	P37892	lophius ame

ALIGNMENTS

RESULT	ID	BLRO_MYRVE	STANDARD	PRT	572 AA.
AC	012737				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	Bilirubin oxidase precursor (EC 1.3.3.5).				
OS	Myrothecium verrucaria.				
OC	Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Myrothecium.				
OX	NCBI_Taxid=5532;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	STRAIN-MT-1;				
RX	MEDLINE=93366794; Pubmed=8360171;				
RA	Kolkeda S., Ando K., Kaji H., Inoue T., Murao S., Takeuchi K.,				
RA	Samejima T.;				
RT	"Molecular cloning of the gene for bilirubin oxidase from Myrothecium				
RT	verrucaria and its expression in yeast.";				
RL	J. Biol. Chem. 268:18801-18809(1993).				
CC	- FUNCTION: OXIDATION OF BILIRUBIN + O(2) = bilirubin + H(2)O.				
CC	- CATALYTIC ACTIVITY: Bilirubin + O(2) = bilirubin + H(2)O.				
CC	- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH				
CC	CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2				
CC	OR NORMAL, AND TYPE 3 OR COUPLED BINDLEAR. CONTAINS 2 BLUE COPPER				
CC	ATOMS PER MOLECULE.				
CC	- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.				
CC	- SIMILARITY: CONTAINS 2 PLASTOCYANIN-LIKE DOMAINS.				
CC	-----				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; D14081; BAA03166.1; -				
DR	EMBL; D12579; BAA02123.1; -				
DR	InterPro: IPR001117; Cu-oxidase.				
DR	Pfam: PF00394; Cu-oxidase.1.				
KW	Signal; Copper; Metal-binding; Oxidoreductase; Glycoprotein; Repeat.				
FT	SIGNAL	1	19		
FT	PROPEP	20	38		
FT	DOMAIN	39	572		
FT	CHAIN	98	194		
FT	DOMAIN	404	526		
FT	METAL	132	132		
FT	METAL	134	134		
FT	METAL	134	134		
FT	METAL	172	172		
FT	METAL	174	174		
FT	METAL	436	436		
FT	METAL	439	439		
FT	METAL	441	441		
FT	METAL	494	494		
FT	METAL	495	495		
FT	METAL	495	495		

FT METAL 496 496 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 500 500 COPPER (TYPE 1) (BY SIMILARITY).
 FT METAL 503 503 COPPER (TYPE 1) (BY SIMILARITY).
 FT CARBOHYD 510 510 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 520 520 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 572 AA: 63947 MW: 584206413035E5EFF CRC64;

Query Match 64.5%; Score 2010; DB 1; Length 572;
 Best Local Similarity 65.4%; Pred. No. 6.8e-135;
 Matches 303; Conservative 71; Mismatches 110; Indels 22; Gaps 10;

QY 1 MISQAIGAVAGLAVIGSSVDARSAGRSTDMPSG-LTRKQTLQSPPLALVEPLPPIP 59
 1 MFKHTLGAALSL-LFNSNNAQASVP--ETSPATGLHFRVAQISQYMFVPLPIP 57
 DB 60 LKAPN-TVPRPNTGEDIYYEMEIRPFSHOITYDLEPANVGDGMSPGTITIVPRGTS 118
 58 VKQPRLVTPNVAGQETWYVEIKPFTHQVYDLSADLVGDGMSPGTIPVPRVET 117
 QY 119 VYFVNSGENTSPNSVHLHGSFRAFPDGAEDTQGEKYDYYPYRQARMLWYDHA 178
 118 VYFVNSGENTSPNSVHLHGSFRAFPDGAEDTQGEKYDYYPYRQARMLWYDHA 175
 DB 118 VYFVNSGENTSPNSVHLHGSFRAFPDGAEDTQGEKYDYYPYRQARMLWYDHA 175
 179 MSTANAYWGAQGYMIDPAEDALNPSGCEPDIPLYLAKRYNADGTLFSTNGEVS 228
 176 MHTAENAYWGAQGYMIDPAEDALNPSGCEPDIPLYLAKRYNADGTLFSTNGEVS 225
 QY 239 SFMGDIYVNGQWPMPLNQPFRKRYRFLNAVSRSPALVATSESETPLPPOVIAADG 298
 236 SFMGDIYVNGQWPMPLNQPFRKRYRFLNAVSRSPALVATSESETPLPPOVIAADG 295
 DB 299 GLEGEVDITDLYISMAERWEVIDSTFAGQSIDIRNLPGA-DGLVEPEFNDTKVMR 357
 296 GLEGEVDITDLYISMAERWEVIDSTFAGQSIDIRNLPGA-DGLVEPEFNDTKVMR 355
 QY 358 FVYDVELESPTSEVYANLADVPFEGGNMPNDPDD--ETTFGRANQMTINVTFS 415
 356 FVYDVELESPTSEVYANLADVPFEGGNMPNDPDD--ETTFGRANQMTINVTFS 407
 DB 416 DVENRLRNPRDVEIWRLENNNGMTHPVHILVDFRVLRSST--ARGVPEYEAAGL 472
 408 DVONRLRNPRDVEIWRLENNNGMTHPVHILVDFRVLRSST--ARGVPEYEAAGL 466
 QY 473 KDYVWLARREVVYEAHYAPFPGVYMLHCHNLIHEDHMAAFNVYLGDYGYNTFEID 532
 467 KDYVWLARREVVYEAHYAPFPGVYMLHCHNLIHEDHMAAFNVYLGDYGYNTFEID 526
 DB 533 PMEELMRPPELIGFENGSGDFSELATDRIOEMASFNRYAADD 578
 527 PMEELMRPPELIGFENGSGDFSELATDRIOEMASFNRYAADD 572

RESULT 2
 COTA_BACSU STRAND: PRT; 513 AA.
 AC P07788; 024818; (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Spore coat protein A.
 GN COTA OR PIG.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124186; PubMed=8969499;
 RA Borris R., Portollik S., Schroeter R.;
 RT "The 52 degrees-55 degrees segment of the Bacillus subtilis
 chromosome: a region devoted to purine uptake and metabolism, and

RT containing the genes cota, gabp and guaA and the pur gene cluster
 within a 34960 bp nucleotide sequence.";
 RT Microbiology 142:3027-3031(1996).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=168 / MARBURG;
 RX MEDLINE=98116660; PubMed=9455482;
 RA Kasahara Y., Nakai S., Ogasawara N., Yata K., Sadaie Y.;
 RT "Sequence analysis of the groES-cota region of the Bacillus subtilis
 genome, containing the restriction/modification system genes.";
 RL DNA Res. 4:335-339(1997).
 RN [3]
 RP SEQUENCE OF 1-37 FROM N.A.
 RX MEDLINE=88011308; PubMed=2821284;
 RA Donovan W., Zheng L., Sandman K., Losick R.;
 RT "Genes encoding spore coat polypeptides from Bacillus subtilis.";
 RL J. Mol. Biol. 196:11-10(1987).
 RN [4]
 RP SEQUENCE OF 1-34 FROM N.A.
 RX MEDLINE=88286730; PubMed=3135411;
 RA Sandman K., Kroos L., Cutting S.M., Youngman P., Losick R.;
 RT "Identification of the promoter for a spore coat protein gene in
 Bacillus subtilis and studies on the regulation of its induction at a
 late stage of sporulation.";
 RL J. Mol. Biol. 200:461-473(1988).
 RN [5]
 RP SEQUENCE OF 1-10 FROM N.A.
 RC STRAIN=168;
 RA Wray L.V., Fisher S.H.;
 RT Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN BROWN PIGMENTATION DURING SPOROGENESIS.
 CC -1- SIMILARITY: TO S.ANTIBIOTICUS PHENOXAZINONE SYNTHASE (PHSN).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: U51115; AAB62305.1; -
 DR EMBL: AB007638; BAA22774.1; ALT_INIT.
 DR EMBL: Z59107; CAB12449.1; -
 DR EMBL: X05678; CAA29165.1; ALT_INIT.
 DR EMBL: X05678; CAA29165.1; -
 DR EMBL: U31756; AAC44642.1; -
 DR PIR: A27393; A27393.
 DR Subtilisin, Bg10490; COTA.
 DR InterPro: IPR001117; Cu-oxidase.
 DR Pfam: PF00394; Cu-oxidase; 1.
 KW Sporulation; Complete proteome.
 FT CONFLICT 347 367 DESRRPKYLASYPSVOHERIQ -> TKAERSRSTPHLTRYS
 FT CONFLICT 451 458 PRGRTPP -> RHAETIL (IN REF. 1).
 FT CONFLICT 451 458 GPVAPPPP -> VRCRAA (IN REF. 1).
 SQ SEQUENCE 513 AA: 58499 MW: 8368838458075F87 CRC64;

Query Match 23.5%; Score 731.5; DB 1; Length 513;
 Best Local Similarity 34.0%; Pred. No. 1.8e-44;
 Matches 180; Conservative 78; Mismatches 178; Indels 93; Gaps 16;
 QY 57 IPLKAPNTVPNNTGEDIYYEMEIRPFSHOITYDLEPANVGDGMSPGTITIVPRGT 116
 7 VDALPPIPTLAKPYQSKERTYIVYMECHTQHLRDLPTRLMGVINGLPPGPIIEVKRNE 66
 DB 117 ESYVRFVNSGENTS-----PNSVHLHGSFRAFPDGMAE-----D 151
 67 NVYVKMMNNLPSTHPLPDHTIHSDSQHEEPVKTIVHLHGCVTPDDSDGYEAMFSKD 126
 QY 152 TTQPGCY--KDIYYNRQAAARMWYHDAASITANAYWGAQGYMIDPAEDALNPS 208
 152 TTQPGCY--KDIYYNRQAAARMWYHDAASITANAYWGAQGYMIDPAEDALNPS 208

```
Db 127 FEOTGRFKEVYHYHPNOQCALMWHDMALTRLVYAGLVANIIDPKERKLPS 186
Cc 209 GYGEFDRLVLTAKRYNADSTLSTNGE-----VSSEFWGDIQVNGCPMPMLN 256
Cc 187 --DEVDPVLLITDRITINEDGSLFPYAPENPSPSLPAPFCGCTILVNKVMYILE 244
Cc 257 VORRKRFRRLNAVSRFALYATSDSTRLPFOVIAADGGLLEGSPVDTLLYSMAE 316
Cc 245 VEPKRYFRVYNASMTRYMLSLDNGD-----FIOIGSDGLLIPRSVKLNSFLAPAE 298
Cc 317 RMFEVDFSEFAGOSIDIRNLPGADGAGVPEFNDMDKVARFVYDVELESDSEVPANL 376
Cc 299 KYIILIDFTAYESESITLANSACCG-DVNPETD--ANIMQFVRYPLAKODSKRKPYL 355
Cc 377 RDVPFEGGNMDEPAPN-----TDDETFTEGRA-----NGQMTNGVTFSDVENRLRNVP 426
Cc 356 ASVPSYQNERIQNIIRLKLKAGTODE---YORPVLLNNKRW-----HDPYTERP 401
Cc 427 R-DTVEIMRLNNSGWTBPHVHILVDFVLSR---STARGV-----PYEA 469
Cc 402 KVGTETIWSITINPTRG-TPHIHLVSRFVLDHRPFDIAVQSGELSTGPAVPPPE 460
Cc 470 AGKDDVWMLARREVVYVEAHYAPFPGVYMLHCNLIHEDHDMMAAFNV 518
Cc 461 KGWKDPTIOAHAGEVLRATFPGYSGRYWCHILEHEDYDMRPMDDIT 509
```

RESULT 3

PHSA_STRAT PHSA_STRAT STANDARD: PRT: 642 AA.

AC 053692; 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Phenoxazinone synthase (EC 1.-.-.) (PHS).

GN PHSA.

OS Streptomycetes antibioticus.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.

OX NCBI_TaxID=1890;

RN [1] SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.

RP STRAIN=IMRO 3720; MEDLINE=96011355; PubMed=7592317;

RC MEDLINE=96011355; PubMed=7592317;

RA Hsieh C.-J., Jones G.H.;

RT "Phenoxazinone synthase from Streptomycetes antibiotics: purification of the large and small enzyme forms."

RT Arch. Biochem. Biophys. 211:55-65(1981).

CC -1- FUNCTION: CATALYZES THE LAST BUT TWO STEPS IN THE PUTATIVE BIOSYNTHETIC PATHWAY OF ACTINOMYCIN.

CC -1- CATALYTIC ACTIVITY: 4.4-methyl-3-hydroxyanthraniloyl pentapeptide + 3 O(2) -> 2 actinomycin acid + 6 H(2)O.

CC -1- COPFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).

CC -1- SUBUNIT: HOMODIMER (SMALL FORM) OR HOMOHETEROMER (LARGE FORM).

CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.

CC -1- SIMILARITY: CONTAINS 2 PLASTOCYANIN-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```
Cc or send an email to license@isb-sib.ch.
Cc -----
Cc EMBL: U04283; AAA86668.1; ALT INIT.
Cc DR InterPro: IPR001117; Cu-oxidase.
Cc DR InterPro: IPR002355; Multicopper oxidase2.
Cc DR Pfam: PF00394; Cu-oxidase; 1.
Cc DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
Cc DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
Cc KW oxidoreductase; Repeat; Metal-binding; Copper;
Cc Antibiotic biosynthesis.
Cc FT INIT MET 0
Cc FT DOMAIN 86 223 0 PLASTOCYANIN-LIKE 1.
Cc FT 223 493 620 PLASTOCYANIN-LIKE 2.
Cc FT METAL 161 163 163 COPPER (TYPE 2) (BY SIMILARITY).
Cc FT METAL 163 163 163 COPPER (TYPE 3) (BY SIMILARITY).
Cc FT METAL 201 201 201 COPPER (TYPE 3) (BY SIMILARITY).
Cc FT METAL 203 203 203 COPPER (TYPE 3) (BY SIMILARITY).
Cc FT METAL 524 524 524 COPPER (TYPE 1) (BY SIMILARITY).
Cc FT METAL 527 527 527 COPPER (TYPE 2) (BY SIMILARITY).
Cc FT METAL 529 529 529 COPPER (TYPE 3) (BY SIMILARITY).
Cc FT METAL 602 602 602 COPPER (TYPE 3) (BY SIMILARITY).
Cc FT METAL 603 603 603 COPPER (TYPE 1) (BY SIMILARITY).
Cc FT METAL 604 604 604 COPPER (TYPE 3) (BY SIMILARITY).
Cc FT METAL 608 608 608 COPPER (TYPE 1) (BY SIMILARITY).
Cc FT METAL 613 613 613 COPPER (TYPE 1) (BY SIMILARITY).
Cc SO SEQUENCE 642 AA; 70113 MW; E0B39C0BA3364E48 CRC64;
```

Query Match 17.2%; Score 536.5; DB 1; Length 642;
Best Local Similarity 29.5%; Pred. No. 1.6e-30;
Matches 184; Conservative 65; Mismatches 200; Indels 175; Gaps 26;

```
Cc 37 LTRKQQLSP--LALYVPLRPPLKARNTVNPNTGEDIYLEMEIRPFSQIIPDE 94
Cc 28 LAKERQAPAPGELPFPALVTP-----VLPRASDEVTRRETELALRPVWRILPOLP 81
Cc 95 PANNVGYDGMSPGPTTV-----PRGESVVRFY-----NSGENTSP- 131
Cc 82 PLTMMGYDGOVPEPTIEVRGCGVRITAMNRIKSGEYVTSVEVGLPGPGPAPRTEBG 141
Cc 132 -----NSVHLGSEFSRADPDGVAEDTTPGGEYKDYYPNRQAARMLWYH 175
Cc 142 RGGVEPNKDVVALPMSVYHLHGAQTGGGNDGADNAVGFDAQISEYFNDAQTMWYH 201
Cc 176 DHAAMSTAFNAWGGAGVYMIODPADALNPSGYEEDIPVLTAKRNAD-----GT 229
Cc 202 DHAAMITRMWMAAGLYGTLVLRDDEEDALGLDS--GDREIPLIADRNLDJDEGRLNR 259
Cc 230 LF-----STNGEVSS-----FWGDIQVNGCPWMLNVOPRKYFRFLNAAVSRFAL 277
Cc 260 LHKITYVQGSNPEIGKPYSIPEFGYTTVNGRIAPYADVDDGWIYLRVNASMARIYML 319
Cc 278 YLATSDSESTRLPFOVIAADGGLLEGSPVDT-----TLYSMAERWEVYIDFSTFAGS 331
Cc 320 VLIEDDPRVPGVHVGISDGLLPRVPVDPDDTLPLVLSAPAREFDLLVDFALGHR 379
Cc 332 IDIRNL-PEA-----DGLG--VEPEDNDKVARFVYDEVLSPTSEPAUL----- 376
Cc 380 LRLVDPGAPAGTDPDGLGVKYP-----VMEFVRTECEE--DSFALPEVLSCSFRM 432
Cc 377 -RDVP--PEGN--W-----DPAN-----PDDETFPG 401
Cc 433 SHDIPRGHILYITPPTGSGSGHPRIWMAVEEDPADVQVAREGVIOGTGADGRTKYR 492
Cc 402 RANGQMTNGVTFSDVENRLRNVPDIVEIWRLENSNGWTHPVHILVDFVLSRSTA 461
Cc 493 R-----TAATFND--GLGFTIGGTHGQWTFNLIS--PLIHPMHILADQVLCR-- 538
Cc 462 RQVEPYEAGL-----KDYVMLARREVVYVEAHYAPF----- 494
Cc 539 ---DAVDASGFDLALAGTTPYRLDPDTPVPLAPMELGKDVFOVPGDGLRYMGKFDGA 595
Cc 495 -GVYMLHCNLIHEDHDMMAAFNV 517
```

DB 596 YGRFMYTHCHLLEHEDMGMRPFV 619

RESULT 4

YACK_ECOLI STANDARD: PRT: 516 AA.

AC P36649; P75655;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable blue-copper protein yack precursor.

GN YACK OR B0123.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI_TaxId=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / W3110;

RA MEDLINE=94261430; PubMed=8202364;

RA Fujita N., Mori H., Yura T., Ishihama A.;

RT the 2.4-4.1 min (110,917-193,643 bp) region.;

RL Nucleic Acids Res. 22:1637-1639(1994).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RA MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RA "The complete genome sequence of Escherichia coli K-12.;"

RL Science 277:1453-1474(1997).

[3]

RP SEQUENCE OF 29-40.

RC STRAIN=K12 / EMG2;

RA MEDLINE=97443375; PubMed=9298646;

RA Link A.J., Robinson K., Church G.M.;

RT "Comparing the predicted and observed properties of proteins encoded in the genome of Escherichia coli K-12.;"

RL Electrophoresis 18:1259-1313(1997).

[4]

RP IDENTIFICATION BY MASS SPECTROMETRY.

RA MEDLINE=99420866; PubMed=10493123;

RA Fountoulakis M., Takacs M.-F., Berndt P., Langen H., Takacs B.;

RT "Enrichment of low abundance proteins of Escherichia coli by hydroxyapatite chromatography.;"

RL Electrophoresis 20:2181-2195(1999).

CC -1- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: periplasmic (potential).

CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.

CC -1- SIMILARITY: CONTRAINS 3 PLASTOCYANIN-LIKE DOMAINS.

CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 464.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: D26562; CAB20297.1; ALT_FRAME.

DR EMBL: AE000121; AAC73234.1; -.

DR PIR: S45200; S45200.

DR Ecogene: Egi2318; yack.

DR InterPro: IPR001117; Cu-oxidase.

DR InterPro: IPR002355; MulticCu_oxidase2.

DR Pfam: PF00394; Cu-oxidase; 1.

DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.

KW Periplasmic; Signal; Copper; Metal-binding; Repeat; Oxidoreductase;

KW Complete proteome.

FT SIGNAL 1 28

FT CHAIN 29 516 PROBABLE BLUE-COPPER PROTEIN YACK.

FT DOMAIN 67 163 PLASTOCYANIN-LIKE 1.

FT DOMAIN 164 410 PLASTOCYANIN-LIKE 2.

FT DOMAIN 411 516 PLASTOCYANIN-LIKE 3.

FT METAL 101 101 COPPER (TYPE 2) (BY SIMILARITY).

FT METAL 103 103 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 141 141 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 143 143 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 443 443 COPPER (TYPE 1) (BY SIMILARITY).

FT METAL 446 446 COPPER (TYPE 2) (BY SIMILARITY).

FT METAL 448 448 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 489 489 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 500 500 COPPER (TYPE 1) (BY SIMILARITY).

FT METAL 501 501 COPPER (TYPE 3) (BY SIMILARITY).

FT METAL 505 505 COPPER (TYPE 1) (BY SIMILARITY).

FT METAL 510 510 COPPER (TYPE 1) (BY SIMILARITY).

SO SEQUENCE 516 AA; 56556 MW; 37D96B1C31CF30B CRC64;

Query Match 13.2%; Score 412.5; DB 1; Length 516;

Best Local Similarity 29.7%; Pred. No. 6,9e-22;

Matches 141; Conservative 55; Mismatches 196; Indels 83; Gaps 19;

QY 100 GYGSGSPGPIIVIPRGTESVVRVNGSEMTSPNSVHLSFSAPRPDGAEDTTPGEEK 159

DB 68 GINGNLGPVVKIQGRKAVTVDIYN--QLTEETLHWGLVPEVDGGPQITPPGKR 125

QY 160 DYYPNROARMLWYHDHMSITENAYMGQGVNVIQPAEDALNLPSCGEFDPILV 219

DB 126 SVTLNDOPATCFWPHPHQGTGRQVAMGLAVIEDEILKMLPKQMGIDVPVILV 185

QY 220 TAKRYADGLT---ESTNGEVSSFMGDIQVNGQPPMLNVPQRK--RRFLNAASRSF 275

DB 186 QDKRFASDQIDYQLDVMTAAVGMFQDITLTNGAIPQ--HAARGLRLRLNGCNARS- 243

QY 276 ALYATSESESTRLPQVIAADGGLLEGVPDPTLISMAEREVVIDESTFAGQSIDR 335

DB 244 -LNFATSDNR---PLVYIASDGLLPPEVKYSELVYLMGEREVLEVND--NKFFDLV 296

QY 336 NLPGAD-GLGVEPEFDNTKVMKRVVDEVLES---PDT-SEVPANLDVFFPEG----- 384

DB 297 TLPSVQMGAIAP-FDKPRPVKRIQPIAISASGALPDTLSIPA---LPSLEGLTVRKL 351

QY 385 -----GNMDPANTDDEFTTFRANG 405

DB 352 QLSMDPMLDMGMQMLMERYGDQAMAGMDHSQMGHMGHGNMNMN--HGKRFDFHNN- 408

QY 406 QMTINVTFSDEVNRLRLRVPRDTVEIWRLENNNSNGWTPTVNHILVDFVLSRTARGVE 465

DB 409 --KINGQARD--MKNKMFPAAKQGYERWVYISGVGMMLHPRFHIGQFRLISN---GAP 461

QY 466 P-YEAGLDVYWLARR--EYVYVEAHYAPFPFGVYVLMHCHNLHEDHMAAFEN 517

DB 462 PAHNRAGMKDYVKEGVNVESEVLKFNHDAPKEHAUYAACHLLEHEDTGMMLGFTV 516

RESULT 5

SUFL_ECOLI STANDARD: PRT: 470 AA.

AC P26648;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Protein sufl precursor.

GN SUFL OR B3017.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

NCBI_TaxID=562;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1234-1238 (1997).
 [2]
 SEQUENCE OF 1-89 FROM N.A.
 RX MEDLINE=92212294; PubMed=1557036;
 RA Coleman J.;
 RT "Characterization of the *Escherichia coli* gene for
 RT 1-acyl-sn-glycerol-3-phosphate acyltransferase (plsc).";
 RL Mol. Gen. Genet. 232:295-303 (1992).
 [3]
 SEQUENCE OF 28-39.
 RC STRAIN-K12 / EMG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 RT in the genome of *Escherichia coli* K-12.";
 RL Electrophoresis 18:1259-1313 (1997).
 CC -1- FUNCTION: INVOLVED IN CELL DIVISION, SUPPRESSES A FTSI MUTATION.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: U28377; AAA69185.1;
 DR EMBL: AE000384; AAC76053.1;
 DR EMBL: M63491; AAA24398.1;
 DR PIR: S20461; S20461.
 DR EcoGene: EGI1376; SUFI.
 DR InterPro: IPR001117; Cu-oxidase.
 DR Pfam: PF00394; Cu-oxidase; 1.
 DR KMW Periplasmic; Signal; Complete proteome.
 FT CHAIN 1 27
 FT DOMAIN 28 470 PROTEIN SUFI.
 FT FT 164 PLASTOCYANIN-LIKE.
 FT SEQUENCE 470 AA; 51858 MW; C843A5A4CB146688 CRC64;

Query Match 10.2%; Score 319; DB 1; Length 470;
 Best Local Similarity 25.3%; Pred. No. 2.5e-15;
 Matches 131; Conservative 67; Mismatches 205; Indels 114; Gaps 21;

QY 10 ALGLAVIGSSVDARSVAGRSTDMPSGLTKRQQLSPPLALYEVPPIPLKAPNTVNP 69
 DB 11 ASGIALCAGAVPLKASAGQ-----QQPLPVPPL-----ES-42
 QY 70 NNEEDILYEMELRPSHQIYPLLEPANNVGYOMSGPITIVPRGESVVRVNSCENT 129
 DB 43 RROQPLF--MTYOR-AHMSFTPGTRASVINGIRYLGPIRWKGD--VKLYSNRLT 96
 QY 130 SPNSVHLHGSFSAPEFGMAEDTTPCEYKDYVPNROAARMLWYHDHAMSITAENYMG 189
 DB 97 ENVSMTYAGIQVCEPLMGARMAKSPADAPVLPTRONATATMYHANPTNRTAQOYVNG 156
 QY 190 QAGVYMIQDAEDALNLPSCYGEFDPVLVTARKYNADGTLFTNGEVSFMDVQVNG 249
 DB 157 LAGMWLVEDEVSKSLPIPNHYGVDDFVILIQDKRLDNFGIRPEVNEPSSGFGVDITLVNG 216
 QY 250 QPAPMLNVQPRKTRFRFLNAVAASRSPALYLTASEDETRLPFOVIAADGGLLEGVDIT 309

DB 217 VOSPYEVSRGWRLRLNLSNSRRYOLQNMNDR-----PLHVISDQGFLLPAPSVKQ 270
 QY 310 LYSMARREVEYIDFST-----FAGQSDIRNL-----PGADGL 343
 DB 271 LSLAPGERRELLVDMSNGDEVSTICGEASIVDRINGCFEFSSTLVSTLVLP-----T 326
 QY 344 GVEPEFNTKRV-NRFVAVDEVLESPTSEVPANLRDVPPEGGWMDPANPTDDETFEGR 402
 DB 327 GLPLV--TDSLPRKLLPTEIMAGS-----PISRDISLGD---DYG----- 363
 QY 403 ANQO-WTINGVTSDEVNRLLRNVRDTEIWRLENNSGWTHPVIHLVDFRYSRSTA 461
 DB 364 INGLMDVNRK-----DY-----TAQGTWERWYTRADE---PQAFHLEGVWFQ---RVN 408
 QY 462 RGVPEY-EEAGLADVVL---AREVYVENHVAFP 494
 DB 409 NGAMPFEDRGMDTVWVDQVELLVYFGQPSMAHFP 445

RESULT 6
 SUFI_SALTY STANDARD; PRT; 470 AA.
 ID SUFI_SALTY
 AC P40799;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Protein sufi precursor
 GN SUFI OR STM3172 OR STY3349.
 OS *Salmonella typhimurium*, and
 OS *Salmonella typhimurium*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Salmonella*.
 CC NCBI_TaxID=602, 601;
 CC
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=LT2;
 RA Cong J., Schmid M.B.;
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
 [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGGC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856 (2001).
 [3]
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mangell K.L., Bentley S.D., Holden M.T.G., Sebalia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Doid L., White N., Barrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holtrova S., Jagels K.,
 RA Krogan A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*
 RT enterica serovar Typhimurium CT18.";
 RL Nature 413:848-852 (2001).
 CC -1- FUNCTION: INVOLVED IN CELL DIVISION, SUPPRESSES A FTSI MUTATION
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 413.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U09309; AAA56679.1; ALT FRAME.
DR EMBL; AE008845; AA122046.1; -
DR EMBL; AL627277; CAD03004.1; -
DR StyGene; SG10522; sufi.
DR InterPro; IPR001117; Cu-oxidase.
DR Pfam; PF00394; Cu-oxidase; 1.
KW Periplasmic; signal; Complete proteome.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 470 PROTEIN SUFI.
FT DOMAIN 68 164 PLASTOCYANIN-LIKE.
SQ SEQUENCE 470 AA: 51858 MW: 61884Da D42B025FB CRC64;

Query Match 10.0% Score 310; DB 1; Length 470;
Best Local Similarity 24.4%; Pred. No. 1,1e-14;
Matches 126; Conservative 70; Mismatches 207; Indels 114; Gaps 20;

10 ALGLAVIGSSVDARSAGSTDMPSGLTKRQTSLPPLALYEVLPPIPKAPNTVPNP 69
11 ASGIALCAGAIPLRANAAGQ-----QQPLVPVPL-----ES 42
DB
70 NNGEDILYEMERPPSHQIYDLEPANVGYDGMSPPTIYVPROTESVYRVNGENT 129
43 RRGQPLF---MQLOR-AHMSFTQGRAPYGVNGRYLGTIRWKGDD--VKLITSNRLA 96
QY 130 SPNSVYLHGSFSRAPPDGAEDTQGEYKDYYPNRQAARMLMYDHAMSTIAENAYNG 189
DB 97 ENVSMTVAGLLPGLPMGPARMSPNADMAPLPTRQSAITIMYANTPNRFAQVYNG 156
QY 190 QGVYVMIQDPAEDALNLPSCYGEFDIPLVLTAKRYNADGTLFSTNGEVSFWGDVYQNG 249
DB 157 LAGMWLVEDDISKTLPTPHNYGVDPFVILQDKRLDNFGPPEVSESGGFGVDDTLVNG 216
QY 250 QVWPMNVQPRKRYRFLNAVSRFALLATSEDETRLPQVYIADGGLGEPDITDT 309
DB 217 AOSPVEVSRGVWRLRLNLSNSRYQLQMSDGR-----ALHVISGDGCFPLAPVSVKQ 270
QY 310 LYISMAERREVVDFST-----FAGOSIDIRNL-----PGADGL 343
DB 271 LSLAPGERREILLVDMNGDEVITCGEASIVDRINGFEPPSSILVSTLVLRP-----T 326
QY 344 GVEPEFNTDKV-MRFVDEVLESPTSEVPANLRDVPPEGGNMDPANPTDEFTFGR 402
DB 327 GLPLPLV--TDNLPMLRLPTEIM-----SGAPVSRDISLGD-----DPG----- 363
QY 403 ANGO-WTNGVYFSDVENLLRNVRDYEIRMLENNSGMTHPVIIHLVDFRVLRSRTA 461
DB 364 INGQIMDVNRIDIT-----AQQGTWERMTVRADM-----PQSHLEGVSFLIRN-V 408
QY 462 RGEVEY-EEAGLKDYYVL---ARREVVYEAHYAPP 494
DB 409 NGAMPEPDRGKMDYVWGVQVELLYVGQPSMPHP 445

RESULT 7
FET5_YEAST
ID FET5_YEAST STANDARD: PRT: 622 AA.
AC P43561;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Iron transport multicopper oxidase FET5 precursor (EC 1.10.1.1).
GN FET5 OR YFL041W.
OS Saccharomyces cerevisiae (baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasakuma S.-I., Sasakuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae."
RL Nat. Genet. 10:261-268(1995).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=98075360; PubMed=9413439;
RA Spizzo T., Byersdorfer C., Duesterhoeft S., Eide D.;
RT "The yeast FET5 gene encodes a FET3-related multicopper oxidase
RT implicated in iron transport."
RL Mol. Gen. Genet. 256:547-556(1997).
CC -1- FUNCTION: IRON TRANSPORT MULTICOPPER OXIDASE, WHICH IS REQUIRED
CC FOR FERROUS IRON HIGH AFFINITY UPTAKE. MAY BE REQUIRED TO OXIDIZE
CC FE(II) AND RELEASE IT FROM THE TRANSPORTER. ESSENTIAL COMPONENT OF
CC COPPER-DEPENDENT IRON TRANSPORT.
CC -1- CORFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH
CC CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE
CC 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MULTICOPPER OXIDASE FAMILY.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50617; BAA09199.1; -
DR EMBL; D44598; BAA08024.1; -
DR HSP; P37064; IAO2.
DR SGD; S0001853; FET5.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR002355; Multicu_oxidase2.
DR Pfam; PF00394; Cu-oxidase; 3.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
KW Glycoprotein; Transmembrane; signal; transport; oxidoreductase;
KW Copper; Repeat; Metal-binding; Iron transport.
FT SIGNAL 1 18
FT CHAIN 19 622 IRON TRANSPORT MULTICOPPER OXIDASE FET5.
FT DOMAIN 19 573 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 574 594 POTENTIAL.
FT DOMAIN 595 622 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 43 146 PLASTOCYANIN-LIKE 1.
FT DOMAIN 192 301 PLASTOCYANIN-LIKE 2.
FT DOMAIN 392 514 PLASTOCYANIN-LIKE 3.
FT METAL 79 79 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 81 81 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 128 128 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 130 130 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 418 418 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 421 421 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 423 423 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 496 496 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 497 497 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 498 498 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 502 502 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL ? ? COPPER (TYPE 1) (BY SIMILARITY).
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 622 AA; 70880 MW; 96157A6C86B7FAF CRC64;

Query Match 7.9%; Score 247.5; DB 1; Length 622;
 Best Local Similarity 22.3%; Pred. No. 4.2e-10;
 Matches 130; Conservative 88; Mismatches 187; Indels 177; Gaps 31;

OY 91 PD-LEPANNVYDGMSPGPTIIVRGTSYVRFVNSGENTSPNSVHLGSF-----SRA 143
 DB 34 PDLHKKRMIGFENGWPLPDHVEKSDRYELYLNGFDATATSLRHGLFQNTSLGNOL 93
 OY 144 PRGMAEDTQ-----PGEYKDYYPNQAARMLYHDHASTAENAKYMQAGAYTMQDP 199
 DB 94 QMDG-PSMTQCPVPGQYLYNFVPEQVGTWYAHMGAQYGD-----GMRGAFIITHD 148
 OY 200 AEDALNLPSCGEPDIPLYLT-----AKRYNADGT-----LESTNGE 236
 DB 149 EEPF-----EYDHERVITLSDHYENKTYTKETLSKYNPTGAEPIQNTLPNTN 200
 OY 237 VSSEFGDVIQVNGQPMPLNVQ-RRYRERELAAVSRSFALYLATSESETRLPQVIA 295
 DB 201 VT-----LDFTGETYLFRLNVLNGLFVSQYIL-----EDHE-----MSIVE 237
 OY 296 ADGGLGPDVDTLYISMAERWEVIDFS-----TFA-----GQ 330
 DB 238 VDSVYVK-PNFTDSTYLSAGQMSVILKAKDKMPTRYAMQIMDTMLDVPPELQNL 296
 OY 331 SIDIR---NLPGADGLGVEP-ERD--NTDKVMPFVDEVLSPDTESEVLANLBDVPEEG 384
 DB 297 TIGMRGHSLEPARALNIECCDLDRANDPYLERLLEROLLAYHDQIYDVAVMNLGDS 356
 OY 385 GNV-----DPAIPTDEFTFGSANGQWTINGVTSDEVNRLLRNPRDTVEIMRL 435
 DB 357 VKAFENNTIYVYPKVPPTLTLLTSGKLASDPRIYG--DINAQLKHN--NDIIEV--V 409
 OY 436 ENNSNGWTPVNHILVDFRYSRSTARGV-----PY-BAAGL-----KDYV 436
 DB 410 LNNYDGRHPFHLGHNFQVQKSPGFHDEAYDESEODEMTVPYNESAPLQPPPEPM 469
 OY 477 WLAREVYVE--AHY-----AFPGVYMLHCH-----NLIH 506
 DB 470 ----RRTVVLPSGCHVYLFRRADNPGWYFCHCHVMDHLOGLASVFIETAVLLQEREKLN 525
 OY 507 EDH-DMAAARNVTVLDGYGNTETFD---FNEPLMRPRPF 543
 DB 526 ENYLDICKADIPVVGNAAGHSNDWFDLGLPQPPPLPKGF 567

RESULT 8
 SUPTL_HAEIN STANDARD: PRT: 311 AA.
 AC P44847;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein sufi homolog precursor.
 GN SUFI OR H10733.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus
 NCBI_TaxId=727;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 Keilavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 Rafterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 Rine L.D., Fritchman J.L., Fuhrman J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RT Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus
 influenzae Rd";
 RL Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: Periplasmic (potential).
 CC -1- SIMILARITY: STRONG, TO E.COLI AND S.TYPHIMURIUM SUFI.
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U32756; AAC22390.1; .
 DR TIGR: H10733; .
 KW Periplasmic; Signal; Complete proteome.
 FT CHAIN 1 28 BY SIMILARITY.
 FT SIGNAL 29 311 PROTEIN SUFI HOMOLOG.
 FT DOMAIN 68 164 PLASTOCYANIN-LIKE.
 SO SEQUENCE 311 AA; 34496 MW; 436173F59F1E1772 CRC64;

Query Match 7.9%; Score 247; DB 1; Length 311;
 Best Local Similarity 31.1%; Pred. No. 1.8e-10;
 Matches 96; Conservative 40; Mismatches 129; Indels 44; Gaps 13;

OY 33 MBSGLTKRQ-----TQLS-----PPLALYEVLPPLPKAPNTVNPNTGEDILYEM 80
 DB 1 MPR-LSRQLTKTAISTALSTVAPAPLAAAREKLAVPPL-----IEVRGRPIVLTWQ 53
 OY 81 EIRPFHQIYPOLEPANMV--GYDGMSPGPTIIVRGTSYVRFVNSGENTSPNSVHL- 136
 DB 54 ETN-----YF-LDGSNNVYWGFGNGVYLGPTIKRISGSFAKLN--HNLPQVALS 102
 OY 137 -HGSFSAPEFDGWAEDTQPEYKDYYPNQAARMLYMDHASTAENAKYMQAGAYVM 195
 DB 103 IQGLQASGELFGGAARVLLKKGESMAPVPIEQPAAQWYSATLANSAYQTYGLAGMWL 162
 OY 196 IQDPAEDALNLPSCGEBEDIPVLTAKRYNADG-TLESTNGEVSFSGVDVIQVNGQPMW 254
 DB 163 IEDEQSLKANLPKRYGVDDIPILIQDMEFNNDLQLEFKQN--QPHFVGNRLVNGIEAPY 220
 OY 255 LNVQPKRYRFRFLNAVRSFALYLATSESETRLPFOVIAADGGLLEGPDVDTLYISM 314
 DB 221 LDVARGWIRLRLLNASTLARAYDLRL--DNDQEMLL---IAQDLGLPRAKYSKSVLSP 274
 OY 315 AERREVID 323
 DB 275 GERAEITLVN 283

RESULT 9
 LAC1_CRYPA STANDARD: PRT: 591 AA.
 ID LAC1_CRYPA
 AC 003966;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Laccase precursor (EC 1.10.3.2) (Benzeneol: oxygen oxidoreductase)
 DE (Urushiol oxidase).
 GN LAC-1.
 OS Cryphonectria parasitica (Chesnut blight fungus) (Endothia
 parasitica).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Diaporthales; Valsaceae; Cryphonectria.
 NCBI_TaxId=5116;


```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92314427; PubMed=1535523;
RT Chai G.H., Larson T.G., Nuss D.L.;
RT "Molecular analysis of the laccase gene from the chestnut blight
RT fungus and selective suppression of its expression in an isogenic
RT hypovirulent strain."
RT Mol. Plant Microbe Interact. 5:119-128(1992).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLE).
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzenesemiquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted (potential).
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M73257; AAA3105.1; -
CC EMBL: S38903; AAA09235.1; -
CC HSRP: P37064; IAO2
CC InterPro: IPR001117; Cu-oxidase.
CC InterPro: IPR002355; Multicu_oxidase2.
CC Pfam: PF00394; Cu-oxidase; 3.
CC PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
CC PROSITE: PS00080; MULTICOPPER_OXIDASE3; 1.
CC Oxidoreductase; Signal; Copper; Metal-binding; Lignin degradation;
CC Glycoprotein; Repeat.
CC SIGNAL 1 20
CC FT CHAIN 21 591
CC FT DOMAIN 66 189
CC FT DOMAIN 198 356
CC FT DOMAIN 416 551
CC FT METAL 126 126
CC FT METAL 128 128
CC FT METAL 171 171
CC FT METAL 173 173
CC FT METAL 463 463
CC FT METAL 466 466
CC FT METAL 468 468
CC FT METAL 533 533
CC FT METAL 534 534
CC FT METAL 535 535
CC FT METAL 539 539
CC FT METAL ? ?
CC FT CARBOHYD 121 121
CC FT CARBOHYD 224 224
CC FT CARBOHYD 242 242
CC FT CARBOHYD 265 265
CC FT CARBOHYD 333 333
CC FT CARBOHYD 407 407
CC FT CARBOHYD 425 425
CC SEQUENCE 591 AA; 64696 MW; B2F44CB2AD77701 CRC64;

```

Query Match 7.0%; Score 219.5; DB 1; Length 591;
 Best Local Similarity 23.6%; Pred. No. 3.8e-08;
 Matches 155; Conservative 78; Mismatches 228; Indels 195; Gaps 39;

```

QY 35 SGLTKRQOTLSPPLALYVPLPIPKAPN-----TVPNPTGE 73
DB 10 SGLIASOLSAAPSLTH-----PLEPRQOPNCNTASNRACWISSGYDITTDYEVKPTLTVG 65
QY 74 DILYEMETIRPFSHQIYTPD-LERANVGVGDGMSPGPTIIVPRTESEVVRVNSGEMTSPN 132

```

```

DB 66 -VRQYDLTLTQAEENWLGPPGVVKEEDMLVNGNITLGPVTHAQWG-DTISVTNNLKYNGT 123
QY 133 SVYLIHG-----SFSRAPFQMAEDTTP--GEYKDYVYPRQAAKMLVMDHMSITAENA 186
DB 124 TTHMHQIRQLNTNLDQGVAGITCEPIPRNGSKTYTFFIAHQVGT- WYSH- FSAQYGN- 180
QY 187 YNGQAGVYMIQDPAEDALNPSGVEEDI--PLVITAKRY-NADGTLFSTNGEVSFWG 242
DB 181 --GIYGAQIDGPA-----SLP-----YDIDLGLPLVSDYYKYTADELVYTSQNAIPA-S 228
QY 243 DYIVGVGQRPWMLNVP-----RKRRREFLAASRSFALVLAISSEDE 286
DB 229 DNVLFNG-----TNINPANTTGGQYKKTITLTPGKRRRLRIINTSENNFVSIVGHS--- 280
QY 287 TRLPFQVIAADGGLLEGPPV--TDPLYSMAERMEVVIDFST-----FA 328
DB 281 ----MTVIESD----FVPPDSFTTDSLFPVIGQRYVITIDAQATNTYMMNTFFGGGCG 332
QY 329 GOSID-----IRNLPGADLGVEPEFDNTDKVRFVYDEVLESPP-TSEVPANLRDVPFP 382
DB 333 GKSNNFYPAAIITHYNGA-----SNSHPTNKGVAPADHEDLDLNLVPPVPRSP-T 382
QY 383 EGNMPPAMPPTDEFTTFRANGQWTINVTSDVE-----NRLRNVPRTVDEI 432
DB 383 SGFVAASDNTLDVQLSTTR--KWTINQSTL-DVDWGHPTIYQYINKSTAMPSTDNVWL 438
QY 433 -----WRLENN-----SNGWTHPVHILVDFRILSR-----TARVEPEYEAAGLMD 474
DB 439 VERANQWATWLIENDPTAGNLPHPIHLHGHDFFVLGSSPNVSPVTAQ--TPTTFPS-SD 495
QY 475 VVWL-----ARREYVVE-----AHVAPFQGVWLHCHNLIHEDHDMAAFNVTVLGD 522
DB 496 VSSLNGNNPDIRRVPVLPKGMWLLFAQTTPNGAMWLHCHIMW-----VSAG 543
QY 523 YGVNTEFTDPMPLRPPRPFLIGEFENGSGPESLAIIDRIQEMASFPVQAQD 578
DB 544 LGNTFLE-----QPSAFVAGLNTN--DVNQLN-----SQCKSMNATYIPSKD 582

```

RESULT 10
 FE73.CANAL STANDARD: PRT; 624 AA.
 AC P78591;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Iron transport multicopper oxidase precursor (EC 1.-.-.-).
 GN FE73
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID:5476;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-1161;
 RA Hundt S., Eck R.;
 RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: IRON TRANSPORT MULTICOPPER FERROXIDASE REQUIRED FOR
 CC FERROUS IRON HIGH AFFINITY UPTAKE. REQUIRED TO OXIDIZE FE(II) AND
 CC RELEASE IT FROM THE TRANSPORTER. ESSENTIAL COMPONENT OF COPPER-
 CC DEPENDENT IRON TRANSPORT (BY SIMILARITY).
 CC -1- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH
 CC CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE
 CC 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL SURFACE BOUND
 CC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE MULTICOPPER OXIDASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its


```

FT METAL 132 132 COPPER (TYPE 3) (BY SIMILARITY)
FT METAL 134 134 COPPER (TYPE 3) (BY SIMILARITY)
FT METAL 425 425 COPPER (TYPE 1) (BY SIMILARITY)
FT METAL 428 428 COPPER (TYPE 2) (BY SIMILARITY)
FT METAL 430 430 COPPER (TYPE 3) (BY SIMILARITY)
FT METAL 430 430 COPPER (TYPE 3) (BY SIMILARITY)
FT METAL 480 480 COPPER (TYPE 1) (BY SIMILARITY)
FT METAL 481 481 COPPER (TYPE 1) (BY SIMILARITY)
FT METAL 482 482 COPPER (TYPE 1) (BY SIMILARITY)
FT METAL 486 486 COPPER (TYPE 1) (BY SIMILARITY)
FT METAL ? ? COPPER (TYPE 1) (BY SIMILARITY)
FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 209 209 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 366 366 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL)
FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL)
SQ SEQUENCE 527 AA: 56094 MW: D9597491F1F79625 CRC64:

```

Query Match 6.4%; Score 199.5; DB 1; Length 527;
 Best Local Similarity 23.3%; Pred. No. 8.4e-07;
 Matches 120; Conservative 62; Mismatches 200; Indels 133; Gaps 28;

```

QY 103 GMSPGPTIIVRGESVYVFNSENGTS---PNSVHLHSGFSRAPFDGMAED----- 151
DB 52 GVPKGPRLTNKDKDEQINIDNLITNETMLKSTIIHNGIFQAG--TNMADGAFAVNCQP 109
QY 152 -TQPGKRYKYYNRQAAMLYHDAHMSITENAYMGAGYVMIODPADALNLSGV 210
DB 110 IATNGSLFDYVDD--AGTFYHSLSTQYCD---GLRGLVYVDDPDANASL---Y 160
QY 211 GEPDIPLVLT-AKRYNA-----DGTLESTNGEVSFWG-----DVIOVNCQPV 252
DB 161 DVDDDTTVITLADWYHTAAKLGPAFAGPDSVLINGLRSGSGDGATNLTVITV----- 215
QY 253 PLANVQPRKREFPLNAASRSFALYIATSEDESTRLPFOYIADGGLGSPVDTITLY 312
DB 216 ---TGGRKRRFLVSTISCDPNFTF---SIDGNMTIIV---GGVNHADLDVDSIQI 263
QY 313 SMAERREVVDFSTFAGQSID---IRNLPADAGLGVPEFDNT---DKMAREVVDEVLE 365
DB 264 FAGQKRSFIIN-----ANQSIDNWIRAIPTNG-----TDTTGGVSAIIRYDAEIE 313
QY 366 ---SPDTSEVPANLRD-VPPREGGNDPANPTDET-----FTFGRANGQWITNG 411
DB 314 PTINATTSVILFETDLPVD-----NPAPRGDPQGVGLDAMSLDPSFGSN--FTINN 366
QY 412 VTF-----SDVENKRLRNVP-RDTVEL-----WLENSNGWT 443
DB 367 ETVVPTVPLVLOILSGAODASLLPMSGYTLTSPNSITIEISPIITTDGAL--NAGAP 424
QY 444 HPHHILVDRVLSRSTARGVEPEAGLKDQVWLARREVVYVAHAPPGVYMLCHN 503
DB 425 HPHHLGHTFSV--RSAGSSTFNANPVRDQVSTGNSGDNVITRTTNDGPMFLHCHI 483
QY 504 LIHEDHMAAFNVTVLGDYGYNTFEIDPMEPLM 538
DB 484 DFHLD---AGFAIVFAEDTA--DTASANVPPTAM 512

```

```

DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Laccase 4 precursor (EC 1.10.3.2) (Benzenediol:oxygen oxidoreductase)
DE (Urishiol oxidase) (Diphenol oxidase).
GN LCCL
OS Thanatephorus cucumeris (Black scurf of potato) (Rhizoctonia solani).
OC Eukaryota: Fungi: Basidiomycota: Hymenomycetes: Heterobasidiomycetes:
OC Heterobasidiomycetes: Ceratobasidiales: Ceratobasidiaceae:
OC mitosporic Ceratobasidiaceae: Rhizoctonia.
OX NCBI_Taxid=107832;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=RS22;
RX MEDLINE=9611523; PubMed=8598061;
RA Hahleithner J.A., Xu F., Brown K.M., Brown S.H., Gollightly E.J.,
RA Hahleithner T., Kauppinen S., Pederson A., Schneider P.;
RT "The identification and characterization of four laccases from the
RT plant pathogenic fungus Rhizoctonia solani."
RL Curr. Genet. 29:395-403(1996).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLE). THIS ISOZYME SHOWS OPTIMAL ACTIVITY AT PH 7.
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) -> 4 benzoquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: IN MYCELIA, AT A HIGHER LEVEL THAN LCCL, LCC2
CC AND LCC3.
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 254277; CAA91042.1;
CC DR InterPro: IPR001117; Cu-oxidase.
CC DR InterPro: IPR002355; MulticCu_oxidase2.
CC DR Pfam: PF00394; Cu-oxidase; 3.
CC DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; FALSE_NEG.
CC DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; FALSE_NEG.
CC KW oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;
CC Lignin degradation; Multigene family; Polymorphism.
FT SIGNAL 1 19
FT CHAIN 20 531
FT DOMAIN 23 146
FT DOMAIN 158 315
FT DOMAIN 384 507
FT METAL 83 83
FT METAL 85 85
FT METAL 128 128
FT METAL 130 130
FT METAL 427 427
FT METAL 430 430
FT METAL 432 432
FT METAL 475 475
FT METAL 480 480
FT METAL 481 481
FT METAL 484 484
FT METAL ? ?
FT CARBOHYD 66 66
FT CARBOHYD 109 109
FT CARBOHYD 186 186
FT CARBOHYD 231 231
FT CARBOHYD 280 280
FT CARBOHYD 395 395
FT VARIANT 42 42
FT VARIANT 119 119

```



```

DB 413 GAPHPHGLHGAFAVY-RSAGSTVYNYNDPIFRDVYSTGTPAGADVITFR-TDNPGFW 470
OY 498 MLCCHNLJIEDHDMMAAFVTVLGDYGVNTERIDMEPLMRRLPLGFEENGSDFE 557
DB 471 FLHC-----HIDFLLEGFAVRED----- 491
OY 558 LAITDRIQEMASFNPAQADD 579
DB 492 -----IPDVASANPVQAQMSD 507

RESULT 15
LAC1_PLEOS STANDARD: PRT: 529 AA.
ID LAC1_PLEOS
AC 012729;
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DE Laccase 1 precursor (EC 1.10.3.2) (Benzenediol:oxygen oxidoreductase)
DE (urishiol oxidase) (Diphenol oxidase).
GN POX1.
OS Pleurotus ostreatus (Oyster mushroom) (White-rot fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Pleurotaceae; Pleurotus.
OX NCBI_TaxID=5322;
RN 11]
RP SEQUENCE FROM N.A.
RC STRAIN=FLORIDA; TISSUE=MYCELIIUM;
RX MEDLINE=95314294; PubMed=7793961;
RA Giardina P., Camilo R., Martirani L., Marzullo L., Palmeri G.,
RA Samia G.;
RT "Cloning and sequencing of a laccase gene from the lignin-degrading
RT basidiomycete Pleurotus ostreatus ";
RL Appl. Environ. Microbiol. 61:2408-2413(1995).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLE).
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzoquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: 234847; CA84356.1; -
CC EMBL: 222591; CA840305.1; -
CC HSSP: P37064; IAO2.
CC InterPro: IPR001117; Cu-oxidase.
CC InterPro: IPR002355; Multicu_oxidase2.
CC Pfam: PF00394; Cu-oxidase; 3.
CC PROSITE: PS00079; MULTICOPPER_OXIDASE1; 2.
CC PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
CC OXidoreductase; Signal; Copper; Metal-binding; Lignin degradation;
CC Glycoprotein; Repeat; Multigene family.
CC SIGNAL 1 23
CC CHAIN 1 24 529
CC DOMAIN 25 159 LACCASE 1.
CC DOMAIN 170 312 PLASTOCYANIN-LIKE 1.
CC DOMAIN 360 499 PLASTOCYANIN-LIKE 2.
CC METAL 96 96 PLASTOCYANIN-LIKE 3.
CC METAL 96 96 COPPER (TYPE 2) (BY SIMILARITY).
CC METAL 98 98 COPPER (TYPE 3) (BY SIMILARITY).
CC METAL 141 141 COPPER (TYPE 3) (BY SIMILARITY).
CC METAL 143 143 COPPER (TYPE 3) (BY SIMILARITY).

```

```

FT METAL 425 425 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 428 428 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 430 430 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 481 481 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 482 482 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 483 483 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 487 487 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 487 487 COPPER (TYPE 1) (BY SIMILARITY).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 465 465 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 529 AA; 56580 MW; 40110AFC486882F CRC64;

```

Query Match 6.2%; Score 193; DB 1; Length 529;
 Best Local Similarity 22.8%; Pred. No. 2.4e-06;
 Matches 103; Conservative 58; Mismatches 180; Indels 110; Gaps 23;

```

OY 124 NSEGTSPN-----SVHLHGSFSRAPFDGNAEDTTQGE-----YKD 160
DB 71 NKGDNFQNLVNLQSLDTMLKTSIHWHGFTQSG--STWADGFAVNOQCIASGNSFLYD 128
OY 161 YVYPNBOAARMLWYHDHAMSITAEAMVGOAGYMIQDPADALNLPSCYGEFDPLVLT 220
DB 129 FNVPDQ--AGTFWYHSHLSTQYCD---GLRGPIYVDPSPHLSL---YDVNADPTIIT 179
OY 221 AKRY-----NADGTLFSTNGEVSFWGDVYQVNGQP---WPMUNVOPRK-YRF 264
DB 180 LEDWYHVVAPOQNAVLPPTADSTLNGKGRFA-----GGPTSAALAVINVESNRYRF 229
OY 265 RFLNAAVNSFALYLATSEDESETRLPFOVTAADGGLGEPVDITLYISAEWEVYIDF 324
DB 230 RLISMCDPRFTF-----SIDGHS---LQVTEADAVNI-VPIVDSIOIFRGQVSVLN- 280
OY 325 STPAGSID--IRNLPGADGLGVEEPDN--TDKYMRF---VVDYLESPTSEVPANL 376
DB 281 ---ANQTVNMYWIRADPNLIGSTG---FDGINSALIRYAGATEDPPTTSSTPLEET 333
OY 377 RDVPFEGGMDPANPTDDE-----TFTGRANGQNTINCVTSVDENRLRNVPPDTV- 430
DB 334 NLVPLENPGAPGPAVPGAGADININLMAAFDVNTFELTINSPPKAPTAPVLAQLIISGATT 393
OY 431 -----EIMRLENN-----SNGWTHPVHILVDFRLSRSTARGVPEAAGL 472
DB 394 AASLLPSGSIYSLEANKVVEISIPALAVGSRPHFLGHRTFDVI-RSAGSTYTNPDTPAR 452
OY 473 KDVVMLARREVVYEAHY-APPGVYMLHCH 502
DB 453 RDVVNVTGTDANDVNTIRFVTDNPGPWFLLCH 483

```

Search completed: October 3, 2002, 11:02:16
 Job time: 440 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:55:41 ; Search time 55.98 Seconds
(without alignments)
1801.644 Million cell updates/sec

Title: US-09-656-640A-2
3114
Sequence: 1 MISQAGVAVLGLAVIGSS.....IQEMASFPVPAQDDDAEE 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirs:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1839	58.7	602	3	Q9P8C3
2	599	19.2	475	2	Q93M03
3	501.5	16.1	637	10	Q9FTS3
4	489.5	15.7	568	10	Q23123
5	476.5	15.3	582	10	Q9AMU4
6	455.5	14.6	614	10	Q9FTS5
7	448	14.4	588	10	Q9FTS6
8	443	14.2	581	10	Q949X8
9	442	14.2	527	16	Q67206
10	431	13.8	591	10	Q9C9A4
11	402.5	12.9	536	2	Q938E6
12	401.5	12.9	502	16	Q9CFW8
13	373.5	12.0	515	16	Q9CJ06
14	288.5	9.3	1662	2	P71431
15	287.5	9.2	513	16	Q9PME8
16	270.5	8.7	468	16	Q9CP61

17	259.5	8.3	463	16	Q9HXM7	Q9hxm7 pseudomonas
18	256	8.2	500	16	Q9KB49	Q9kb49 bacillus ha
19	247.5	7.9	460	2	Q9X3V2	Q9x3v2 pseudomonas
20	246	7.9	504	16	Q53858	Q53858 mycobacteri
21	245	7.9	675	2	Q9KJ08	Q9kj08 marionomas
22	237.5	7.6	561	3	Q96WN0	Q96wn0 botrytis ci
23	236.5	7.5	533	3	Q60199	Q60199 pleurotus o
24	232.5	7.5	464	16	Q9ZQV6	Q9zqv6 rhizobium m
25	231	7.4	577	10	Q9RJD5	Q9rjd5 arabidopsis
26	229	7.4	591	10	Q9AUI2	Q9aui2 pinus taeda
27	224.5	7.2	518	3	Q9HDS9	Q9hds9 polytorus c
28	223.5	7.2	609	3	Q9C1Z1	Q9c1z1 pichia past
29	222	7.1	508	3	Q9UVQ2	Q9uvq2 pycnoporus
30	221	7.1	351	2	Q93F47	Q93f47 pseudomonas
31	220	7.1	516	3	Q9Y782	Q9y782 coprinus ci
32	218.5	7.0	567	10	Q941X2	Q941x2 oryza sativ
33	218.5	7.0	589	3	Q9C497	Q9c497 glomerella
34	218.5	7.0	721	16	Q9PA43	Q9pa43 xyella fas
35	217.5	7.0	517	3	Q9UVQ5	Q9uvq5 marasmius q
36	217.5	7.0	586	10	Q9AUI6	Q9aui6 pinus taeda
37	214.5	6.9	517	3	Q12571	Q12571 basidiomyc
38	214	6.9	464	16	Q989B7	Q989b7 rhizobium l
39	213	6.8	635	3	Q96WT3	Q96wt3 candida gla
40	212.5	6.8	569	10	Q9FY79	Q9fy79 arabidopsis
41	209.5	6.7	580	10	Q9ZQW2	Q9zqw2 populus tri
42	208.5	6.7	517	3	Q9HDQ0	Q9hdq0 trameetes tr
43	207	6.6	581	3	Q96WM9	Q96wm9 botrytis ci
44	206.5	6.6	585	10	Q24044	Q24044 litrodendro
45	205.5	6.6	352	2	Q93F57	Q93f57 pseudomonas

ALIGNMENTS

RESULT	ID	Q9P8C3	PRELIMINARY;	PRT;	602 AA.
AC	Q9P8C3;				
DT	01-OCT-2000 (T-EMBLrel. 15, Created)				
DT	01-OCT-2000 (T-EMBLrel. 15, Last sequence update)				
DT	01-DEC-2001 (T-EMBLrel. 19, Last annotation update)				
DE	POLYPHENOL OXIDASE PRECURSOR.				
GN	PPOA.				
OS	Acromonium murtorum.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;				
OC	Hypocreales; Hypocreaceae; mitosporic Hypocreaceae; Acromonium.				
OX	NCBI_TaxId=45278;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPRAIN-CBS 157.72;				
RX	MEDLINE=21268855; PubMed=11375170;				
RA	Gouka R.J., van der Heiden M., Swarthoff T., Verrills C.T.;				
RT	"Cloning of a phenol oxidase gene from Acromonium murtorum and its				
RT	expression in Aspergillus awamori."				
RL	Appl. Environ. Microbiol. 67:2610-2616(2001).				
DR	EMBL: AJ271104; CAB75422.1; .				
KW	Signal.				
FT	SIGNAL.	1	21	POTENTIAL.	
FT	CHAIN	62	602	POLYPHENOL OXIDASE.	
SO	SEQUENCE	602 AA;	66920 MW;	0303D991405228A3 CRC64;	
Query Match					
Best Local Similarity 58.7%; Score 1829; DB 3; Length 602;					
Matches 352; Conservative 83; Mismatches 131; Indels 34; Gaps 7;					
OY	5 AICVAVLGLAVIGSSV-----DARSVAGRSND-----MP5GLTKROTOLS 45				
DB	6 ALRAALALYLSIKRAGQAMPKRELDIPFEANALALAIYEDDPANDLQRRSPANDLQRRSPLS 65				
OY	46 PLALYEVPLIPPLKAP-NTVPNPNTGEDILLYEMEIRPFSHQIYPDLEPANMGYDGM 104				
DB	66 PAVTLQAPLIPVKEPLFTVTNPNVNGCEIDYVEIRKHSQVFPDLGPAVLGVGDI 125				

RESULT	2		
Q93M03			
ID	Q93M03	PRELIMINARY:	PRT: 475 AA.
AC	Q93M03		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	OXIDOREDUCTASE-LIKE PROTEIN.		
GN	AUR21.		
OS	Streptomyces aureofaciens.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
OC	Actinomycetales; Streptomycinae; Streptocetaceae; Streptomyces		
OX	NCBI_TaxID=1894;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CCM3239;		
RA	Korshak J., Bistakova J., Novakova R., Homeroova D., Rezuchova B.		
RT	"Cloning and characterization of a new polyketide gene cluster in		
RT	Streptomyces aureofaciens CCM3239."		
RL	Submitted (May-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AY033994; AAK61713.1.		
DQ	SEQUENCE 475 AA; 52602 MW; A2ED94AC38DD0C9 CRC64;		

[illegible][illegible]


```

OY 125 -SGENTSPNSVHLHGSFSRAPFDG-----MAEDTTPGEXKYKDYYPNRQ 167
      | | | | |
DB 139 IRRGCVPTVYVHLHGSAHHPQSGSAFAMFTAGFGFTGRAMSTPT-----YTPNNQ 190
OY 168 AARMLWYHDHAMSITAEANVMGQAGVYTIODPAEDA-LNIPSGYGEFDLPVLTARKYNA 226
      | | | | |
DB 191 SFGVLYVHHDHALGLTRANLAGLIGAYVIRNPVAPLPG-DEFDRLMLADRSFYA 249
OY 227 DGLTFSTNCEV-----SSPMGVIOVNGQPMPLNVQPKRYPRFLNAVSRPAL 277
      | | | | |
DB 250 DGSITANNYGIIPNIPQOPPEYFGAITYNGKAMPFLAVARRRYFRLLINISNAFFNL 309
OY 278 YIATSEDESETRLPFOVYIADGGLLEGVPDITLYISAEKMEVEYIDESTFAGOSIDIRN- 336
      | | | | |
DB 310 SLTNG-----LPFTVGSOTNYLSKPVTAASLTVSAETFDVVDVDSQTSSEALVNT 363
OY 337 --LPGADGLGVEPEEDNTDKVRFYDEVLESFDTSEVPANLRD-----VP 380
      | | | | |
DB 364 APYPYDGC--QAPNDLNGKVMKVFISPA-KAKDTSRVPAKLIDYVAVAEEAVQRXYV 419
OY 381 FPEGGWMDANPTDDETFEFGANGQWTINGVTFSPVENRLKNVR-DTVEIWRLENNNS 439
      | | | | |
DB 420 MYEYEAATGNPT-----HLYINGKRLDEPAT-----ETFRPGTTEVWEVINLT 463
OY 440 NGMTHPVHILVDFRVLRSSTARGVEPYEA--AGLKDYY-----476
      | | | | |
DB 464 PD-NHPLHLHLATFOA--TRVRGLVDEDAFKGCMAKLNDVAVCNVSRHVGEEVAVPEH 519
OY 477 ---WLARREVVYEAHYA-----PF-----PGVYMLHCHNLHEDHDM 512
      | | | | |
DB 520 EKGW--KNVVKIAPGYMTTYVKKFPMVDSGKRPDPDATAEPG-YVYHCHILHEDNAMI 575

RESULT 6
OQFTS5 PRELIMINARY: PRT; 614 AA.
AC 09FTS5;
DT 01-MAR-2001 (TREMBLrel, 16, Created)
DT 01-MAR-2001 (TREMBLrel, 16, last sequence update)
DT 01-JUN-2001 (TREMBLrel, 17, last annotation update)
DE PUTATIVE SPORE COAT PROTEIN-LIKE PROTEIN.
GN P0409B08.12.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0409B08."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002860; BAB18285.1;
DR InterPro; IPR002355; MultiCu_oxidase2.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; UNKNOWN_1.
KW Coat protein.
SQ SEQUENCE 614 AA; 68321 MW; 80C6CE856E305193 CRC64;

```

```

Query Match 14.6%; Score 455.5; DB 10; Length 614;
Best local similarity 27.6%; Pred. No. 3e-25;
Matches 171; Conservative 75; Mismatches 218; Indels 155; Gaps 30;

```

```

OY 20 SYDARSVAARSTDMPSGLT---KROTOISPLIALYVPLIPPLKAPNTV-----66
      | | | | |
DB 18 SYVSKSRMGWASDDPDYTPAADPSPPTAARPLRSTADLPKYVPLQPMARIQGYGINQEG 77
OY 67 -PAPNTGEDIIVYEMELRFSHQIYDPLEPANNVGYDGMS-----PGPIIYPRGTEESV 119
      | | | | |
DB 78 FVPV-TNLTIGMKN---TWQFHRDMPTPTVFEY-GOSLQATATPEPGPIVAVRNVPLV 130

```

```

OY 120 VRFVNSGENTS-----PNSVHLHGSFSRAPFDGMA-----EDTTPG 156
      | | | | |
DB 131 VYMNENLPRDHLHMPDPTVTAIRPKNGGVPTVYVHLHGAAQAPRSDSHARAWMTFRDAENG 190
OY 157 E---YDYVYYPNRQ-AARMLWYHDHAMSITAEANVMGQAGVYTIODP-AEDALNLEPSYG 211
      | | | | |
DB 191 STWQTKTYTPYVNPQAPAGINWYHDHALGLTRASILAGLAAAYIVWEPELEMPFNLPSS-G 248
OY 212 EPDIPVLTAKRNNAGTLE-STNGEVSS-----FMGDYIOVNGQPMPLNVQPKRY 262
      | | | | |
DB 249 EFDLHLVLDKRNVDGTLFMDTVGAVPSVHPQWOPETGEVITVNGKAMPFOAVQRRXY 308
OY 263 RFRFLNAVSRPALYIATSEDESETRLPFOVYIADGGLLEGVPDITLYISAEKMEVEYI 322
      | | | | |
DB 309 RLRLNANARYLINIRFSNG-----LPFTVYASDATVLSRPVTVSNLLSPAEIRDIVY 362
OY 323 DSTFAGQS---IDIRN--LPGADGLGVEPEEDNTDKVRFY-----DEVLESFD 368
      | | | | |
DB 363 DFLVVPNPNATDIELNSAPYPEPTGPANATLDG--KYMAFNVSAKQWGDMPMQEPE 420
OY 369 TSEVPANLRDVPPEGGNMDPA-----NPTDETFEFGANGQWTINGVTFSDYE 418
      | | | | |
DB 421 NSTVPEET-GVPPAKYTALEPTPKTRYIVLYENMTSNDPTAKTAN--LYINGKRLDEPP 477
OY 419 NRLLRNP-RDTVEIWRLENNNSGWTHPVHILVDFRVLRSSTARGVEPYEAGLK----473
      | | | | |
DB 478 T---ETPISGTELMHWVNLTPD-NHPLHLHLAEFOAVQMLQVDPDPFKSCMLKHNDT 532
OY 474 -----DVMYLAREVVYEAHYA-----PF-----493
      | | | | |
DB 533 FACNLDQHAVALQVPEEKTW--KNVVKIAPRAYVTSVVAERLVHNNMPYFDATAA 589
OY 494 PGVYMLHCHNLHEDHDM 512
      | | | | |
DB 590 PG-YVYHCHILHEDNAMI 607

```

```

RESULT 7
OQFTS6 PRELIMINARY: PRT; 588 AA.
AC 09FTS6;
DT 01-MAR-2001 (TREMBLrel, 16, Created)
DT 01-MAR-2001 (TREMBLrel, 16, last sequence update)
DT 01-JUN-2001 (TREMBLrel, 17, last annotation update)
DE PUTATIVE SPORE COAT PROTEIN-LIKE PROTEIN.
GN P0409B08.11 OR P0044P08.29.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P044F08."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002860; BAB18284.1;
DR EMBL; AP002909; BAB21194.1;
DR InterPro; IPR002355; MultiCu_oxidase2.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; UNKNOWN_1.
KW Coat protein.
SQ SEQUENCE 588 AA; 64290 MW; 6C6CB995060C4A82 CRC64;

```


Best Local Similarity 28.7%; Pred. No. 2,4e-24;
Matches 165; Conservative 77; Mismatches 204; Indels 128; Gaps 29;

QY 10 ALGAVIGGS--SYDANSVAGSRDMSGLKRTQTSPLALYEVLPPIPLAPN--- 64
DB 14 ALGFSVGLLSLSCGGGTTSSGSGSLTKSLNIPGYFLF-----PDQR 62
QY 65 -----TVPNPTMGEDILYXEMEIRPFSSHQIYPDLEPAMNMGVDSGPGPTIIVP 113
DB 63 VSTAKMTTLEVIPIKSTF--DMLYEID-----NEYPVIFLR 98
QY 114 RGTSSVVRPV--NSEGNTPNSVHLHSGFSRAPF--DGAETTPGCE--YKDYYPNP 167
DB 99 KGQTFSDAFVNNSGED--SLIHWG--FRAPKMSDGHPIYAVVDGGEYSYDPDTIIDRS 153
QY 168 AARMANLHDHMASTTAENAYMGAGVMIODPADALN--LPSGYGEDPIPLVTAKRYN 225
DB 154 GT--YFHPHPRHGRGTGYOYVYGLAGMIIEDEDNLKQALDLEYVIDIPLIQQDKTFD 211
QY 226 ADGTL--FSTNGEVSSFWGDVYIQVNGOPMLNVOPRRKYRFRFLNAAVSRFALYATSED 284
DB 212 SSGQLVYNPNGHM--GFWGDTLLVNLTPNPYMDVERKIYFRILNGSNAPRYRLALLRNGQ 270
QY 285 SEFRLLPQVIAADGGLLEGPDITDLYISMAERMEVYIDF--STPAGOSIDIRNLPG--- 339
DB 271 ---RMRFVWIGVEGGLDTPREVNELLVAPGERIDILLVDFPDASVNDVIKLYNPMLIG 327
QY 340 -----ADGLGVEPEF-----DNTD--KVMRFVYDEVLESPTSEVPANLRDV 379
DB 328 MGMTMGMGMGEMRGMGMGMGMNMDGMADNSEFEVMEFRYTK--DNAVKSITQRLSEV 385
QY 380 PPEPGMNDPANPTD--ETFTFGANGQWTINGVTF-----SDVEN--RLLRNVP 428
DB 386 -----TPIN--TDGAQVORITLGMRRMYETTINGETMEGQYANQOINNPVKVLEQNG 436
QY 429 TVEIWRLENNNGTHTVHILHVDERLYLSRTAGVEPYEA--GLKDYVWLAREVYV 486
DB 437 DVYIIEYVNT--GMTHPMHNGFOVLEERS---LGPLRATLDGMDVIVAMETVRI 491
QY 487 E---AHYAPPGVYMLHCHNLIHEDHMAAFV 517
DB 492 AVDMSPHYNEHOIYLHCHITLHDEGMVNYRY 525

RESULT 10
Q9C9A4 PRELIMINARY; PRT: 591 AA.

AC Q9C9A4;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE SPOR COAT PROTEIN-LIKE PROTEIN.
GN F2JN20.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLUMBIA;
RA MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen A., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,

RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
DR EMBL: AC016972; MAG51692.1;
KM Coat protein.
SQ SEQUENCE 591 AA; 67358 MW; B0CE7B0BE0F05AAF CRC64;

Query Match 13.8%; Score 431; DB 10; Length 591;
Best Local Similarity 26.6%; Pred. No. 1,8e-23;
Matches 143; Conservative 75; Mismatches 180; Indels 140; Gaps 22;

QY 88 QIYDLEPANNVGYDMS-----PGPTIIVPGTSESVRFVN----- 124
DB 80 KFHRLD--PAPVFAYGTSKRSATVPPTIEAVYGVDTVYTRNLPLHLPMPTISPA 136
QY 125 -SGENTSPNSVHLHGSFRAPFGMAE-----DTPQGEYDYYPNQAARMLYH 175
DB 139 IPKHGIPYVYHLHGHEPTSDGMADSWETAGFKETGSKWTKTTHYVYKQOPGNMYH 198
QY 176 DHAMSTAEANAAYMGAGVYMIQ--DPADALNLPSCYGEFDPIPLVTAKRYNADGTLF--- 231
DB 199 DHAAGLRLVNLGLAGSLYILRHSSVSPRLRPG--REFDRPLVIFPRSRKDGSIYMA 257
QY 232 -STNGEVSSFW---GDVIOVNGQPMNLNVOPRRKYRFRFLNAAVSRFALYATSEDS 285
DB 258 TGNPRTIHPQMPYFGDAILVNGKAMPRLTVRRKRYFRITNASNARFFRFSNG--- 314
QY 286 ETRLPQVIAADGGLLEGPDITDLYISMAERMEVYIDESTFAGOSIDIRN---LPADG 342
DB 315 ---LDFIVGSDSAYLAKPVSTKSVLLAPSEIYDVLDFSKSTSKYALILNNAPYPG 371
QY 343 LGVEPEFNDTKVMRFVYDEVLESPTSEVPANLRDVPEPGMNDPANPT----- 393
DB 372 --DPVTEENSKVMKFLINKSE--VDTSIIPKKLIEP-----PAHSTSTRTYIA 419
QY 394 -----DDETFEGANGQWTINGVTF-----DVENRLRNVPDTEIWRLENN 438
DB 420 MFEYVSSIDEPHTLY-----INGLPYNAPYTERPKIGTSEKSLMLTMKWEVINL 471
QY 439 SNGTTPVHILHVDERLYLSRTAGVEPY-----EAAGLKDYV----- 476
DB 472 TED--NRPFLHILGLFKYLEOTALVKSEFTTECKTRNDAAKCEISKYARGNKTAIVYHER 530
QY 477 -W-----LAREVYVEAHYA-----PPGVYMLHCHNLIHEDHMAAF 515
DB 531 GWNKVFMPMGHYTKILVRSYIHSNSYSFADTQDEG--YVYCHILHDHNNMMRRF 587

RESULT 11
Q938E6 PRELIMINARY; PRT: 536 AA.

AC Q938E6;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE MULTICOPPER OXIDASE.
GN CUID.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee I.S., Lim S.Y., Choi S.-Y.;
RT "cuid, a gene encoding a multicopper oxidase from S. typhimurium;
cloning and characterization.";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY053392; AAL15149.1; -
SQ SEQUENCE 536 AA; 58532 MW; DAE43663D1B12B74 CRC64;

Query Match	Score	DB 2	Length	536;
12:98;				
Score	402.5;			
DB 2;				
Length	536;			

QY	100	GYDGSQPCPTTIVRGSGSVYRFEVNS - GENMSPNSVILHCSFSAAPPDGAADPTTQGEY	158
Dd	68	GYNGNLGPAVOLVILKKGSAVYDINHQALED - - - TLHHGLETPIGVDSGPGGIIIPAGCT	124
QY	159	KDYUYPNRQAAKMYLMDHDMASITAEANAUYGQAGVYMIODPAEDALMLPSGSGFEFDPLV	218
Dd	125	RYVEFTPEQRAATGCMHPRHKHCKTGQGVAMGLAGVLIEDEDEIKRLAPRQMGIDVDPVI	184
QY	219	LTAKRYNADGTL - - - FSTNGEYVSFFMGDVLYQVNGPMPMLNVQPRKYRFRFLNAAVSRSF	275
Dd	185	IQDRFSSADGQIDYQDLMTAAVWMPFDOTLTLTGAIYPOHSAPKGMWLRRLNCGNARS -	243
QY	276	ALYUATSDESTRLPFOVIADGGLGCPDPTDPLYSMEERMEVYVDEFTPGAGSIDIR	335
Dd	244	-LNTIASDNR - - - PLVYIASDGGELLEBPVKVYELPLMBGREFVVIDISD - - GKARDLV	296
QY	336	NLPGAD - GLAVEPEFDNTDKMYRFVYVDEVLES - - - PDT - SEVPANILRDVFPREG - - -	384
Dd	297	TLVPSQGMMAIAP - FDKPRHPYMRIGQPAITASGTLPTFTTMA - - - LPSLEGILYRNL	351
QY	385	-----GNMDPA - - - - - 390	
Dd	352	KLSMDPRIDMGMQMLKKYGAQAMSGMDHDSNNAHMGNGMGGEMDHGMDHSGMNHG	411
QY	391	---NPTDEFTPEFRANGQWNTINGVPSDVENRLRNVPRDQVIEIMPLENNSGWTHPRH	447
Dd	412	AMGMNHNKGATDFFINAN - - - FTKQVFD - - MNKMPRAQGRHNRWYISGVGMILTRFH	466
QY	448	IHLVDFRVLRSKSTARVEP - YEAAGLRDVYVL - - AREVVUYEAAHYAPFEGVYMLHCNHL	504
Dd	467	IHGQFRLSEN - - - GKAPAAHPRGCKMDQTVRVEGISEVLVKPRDHQAKHNAVMAHCHLL	523
QY	505	IHEHDMMMAFNV 517	
Dd	524	EHEDTGMMLGFTV 536	
RESULT	12		
Q98FW8			
AC	Q98FW8	PRELIMINARY;	PRT; 502 AA.
DT	01-OCT-2001 (TEMBLrel. 18, Created)		
DT	01-OCT-2001 (TEMBLrel. 18, last sequence update)		
DT	01-OCT-2001 (TEMBLrel. 18, last annotation update)		
DE	POTATIVE PERIPLASMIC OXIDOREDUCTASE.		
GN	ML3586.		
OS	Rhizobium loti (Mesorhizobium loti).		
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;		
CC	Phyllobacteriaceae; Mesorhizobium.		
OX	NCBI_TaxID=381;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MAFF30309;		
RX	MEDLINE=21082930; PubMed=11214968;		
RA	Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,		
RA	Metanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,		
RA	Koshida T., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,		
RA	Takeuchi C., Yamada M., Tabata S.;		
RT	*Complete genome structure of the nitrogen-fixing symbiotic bacterium		
RL	Mesorhizobium loti.		
RL	DNA Res. 7:331-338(2000).		
DR	EMBL; AP003002; BAB50448.1; -		
DR	InterPro; IPR00117; Cu-oxidase.		
DR	Pfam; PF00394; Cu-oxidase; 3.		

KW : Complete proteome.
SQ : SEQUENCE 502 AA; 53995 MW; AE96D894241B36CE CRC64;

Query Match	Score	DB	Length
12.98	401.5	16	502

QY	7	GAVLLAGLAVIGSSVDAKRVNG-RSTMP	-----GLTKRTOGLSP-LALXEY	53	
Db	5	GFLAAGLA--LVSERAAQMADDPMDGCGHAGHDGAA	PQGTVAVPSEGEVREL	61	
QY	54	PLPLPKAPNTVNPNTGEDILLYEMEIRFFSHQIY	PDLEPANMVSDDMSGPTTIIV	113	
Db	62	PL-----LANEASRPGFKATLTAGSATARFAGL	-----DPIILANTSGPDLLEAV	110	
QY	114	RGTESVAFVNSGENTSPNSVHLHGSSRAFDGWAED	TTQPGEYKDYYPNROA-ARML	172	
Db	111	EGDRVEITFARILANEA-STIHWGMGVPPADODGN	PMDPATGTDRTYSFDLPLEASGSY	169	
QY	173	MYHDHAMSITFENAYMOCAGVYMIODPAEALNL	PSGYGEFDIPLVLTARKYNAADGLFS	232	
Db	170	WYHPHPHCKTAEQYRGLAGAFVKKPKADP--	TPAYIG--DTVLYPTDLRLAADGTLPD	224	
QY	233	T-----NGEVSSFWGDIVIQNGCPWMLNVO--	PRKRFREFLNAAVSRSEFALYATSEDS	285	
Db	225	NTMTDLMMGRV---GDHVLVNGKNLTLYPFGAK	RFRFNTATNR-----FLRLSPDG	276	
QY	286	ETRLPFOYIADGGLLEGSPVDTDTLYISMAER	EVVIDESTFAGQSIDIRNLPGADGLV	345	
Db	277	AS---MTIIGDGLLEAPVAAGDILLSPAREL	LYVSF-----DKPGAAL-T	321	
QY	346	EPEFDN-----TDKVMRFVDEVLSPDTS	SEVPANLRDVPFEGGWMDP	389	
Db	322	TLDIDRCMGCGRPADAGLTLLTVNVSQTEAD	PVPLDRLRLPTQL-----	368	
QY	390	ANPDDDETFTFG-----RANGQ--	WTINGVTSQDENRLRNVPRTVEIMLENNSG	441	
Db	369	CAPAVSRREFVFTETEMANNASGEMCPLINGA	AF-DMQ-RIDVAKKGOVELMELVNAD-	425	
QY	442	WTHGVHILHLDLFRVLSSTARQVPEPYAGL	KDYVLAAREVYVEHNAVPEFGYWLHC	501	
Db	426	MDPHEFHGHTQFOVYEHHERGNGISPKAY	RAWKDTVNVNARGEAAVLLRL-ODRGP	PRMYHC	484
QY	502	HNLIHEDHDMMAAFNV	517		
Db	485	HILHEQLGMGVVDV	500		
RESULT 13					
09CJ06					
09CJ06		PRELIMINARY;	PRT;	515 AA.	
AC	09CJ06;				
DT	01-JUN-2001 (TremBrel, 17, Created)				
DT	01-JUN-2001 (TremBrel, 17, Last sequence update)				
DT	01-DEC-2001 (TremBrel, 19, last annotation update)				
DE	HYPOTHEtical PROTEIN PM1938.				
GN	PM1938.				
OS	Pasteurella multocida.				
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;				
OC	Pasteurella.				
OX	NCBI_Taxid=747;				
RN	11				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=PW70.				
RX	MEDLINE=21145866; PubMed=11248100;				
RA	May B.J., Zhang Q., Li L.L., Paustian M.L., Whitam T.S., Kapur V.;				
RT	"Complete genomic sequence of Pasteurella multocida PM70."				
RL	Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).				
DB	EMBL: AB006230; AAK04022.1; --				
DR	InterPro: IPR001117; Cu-oxidase.				
DR	InterPro: IPR002355; MultiCu_oxidase2.				
DR	Pfam: PF00394; Cu-oxidase; 2.				

DR PROSITE: PS00080; MULTICOOPER.OXIDASE2; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 515 AA; 58572 MW; 1665D9B6C7C6743F CRC64;

Query Match 12.0%; Score 373.5; DB 16; Length 515;
 Best Local Similarity 26.4%; Pred. No. 2,6e-19;
 Matches 142; Conservative 69; Mismatches 200; Indels 127; Gaps 22;

```

OY 33 MPGLKRTQTSPLALVEPLPIPLKAPNTVPNPNGEDLLYYEMELRPSHOIYD 92
DB 51 MSGLS--LNGILPKLA-----NOSTDAGLFTYTLKAEPIKIR-AD 89
OY 93 LEPANVYDGMSPGPTTIYPRGTSVVRVNSGENTSPNVHLHGSFRAPDGAEDT 152
DB 90 NKTERMAVNGOLPQPIQIEFGDYEIEFIN--HLRPSVIMHGDVNEADGNPMDM 147
OY 153 TDGEYKDYPPNRQ--AARMLYHDAHMSITAEVAGAVYIODPAEDALNDSGYG 211
DB 148 VEPQGRKVVRETLPGSAGTYWYHPRHDSQVYKGLAGTFVYAKNDPLAHLPEOH- 206
OY 212 EPDIPVLAKRYNADGTLST-----NGEVSFWGDVIOVNGOPMMLNVPKRYFR 265
DB 207 ----WISDLRLNADGTLIPANTMLDMNGRE---GEFVLINGYOPOIOVKTNE-RIR 256
OY 266 FLNAVSRSFALYATSESETRLPFOVIAADGLEGP-VDTDTLYISAEEMEVYIDF 324
DB 257 LMAATGARFRLNI-----PEVKMIYVTEGEGLEKPRAPDELLTFAEVEYIM-- 367
OY 325 STPAGOSIDIRNLPGADGLGVEPEFNDTKVMRFVYDEVLESP-----DTSEV 372
DB 308 ---VGETOGTVNL-----QSGYVDRKRM-----VOEQPKDLTLATIQKSEPVQI 350
OY 373 PANTRVPEPEGNMD-----PANTDDEFTFGR 402
DB 351 PESLRSLP-----NMDEPKVOIRSEKMMNTNMPNMHGHHAATTPTDNPPIPM-- 403
OY 403 ANQWITNGVTSDEVNRLRNVRDVTVEIMRLNNSNGVTHILVDFVLSRSTAR 462
DB 404 MNGMFLINGCTFD--KNRIDFVAKLNEVEQWEIFNESH--MDHFFHLHGTFEVIQTLNG 460
OY 463 GVEPYPAAGLKDVYMLAKREVYV---EAHYAPFPGVYMLHCHNLHEDHMAARV 517
DB 461 KTFEPGKALKDYNLRPEYKILIRKQGH---TGKIMHCHILHLENLGMGMKRV 514

RESULT 14
P71431 PRELIMINARY; PRT; 1662 AA.
AC P71431;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MOFA PROTEIN PRECURSOR.
GN MOFA.
OS Lepothrix discophora.
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC Lepothrix.
OX NCBI_Taxid=89;
RN [1]
RN [1] SEQUENCE OF 1-1150 FROM N.A.
RA STRAIN-SS-1;
RC Corstjens P.L.;
RT Identification and Molecular Analysis of the Lepothrix discophora
RT SS-1 mofA Gene, a Gene Putatively Encoding a Manganese Oxidizing
RT Protein with Copper Domains.
RN Geomicrobiol. J. 14:91-108(1997).
RN [3]

```

RP SEQUENCE FROM N.A.
 RC STRAIN-SS-1;
 RA Corstjens P.L.;
 RL Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: Z25774; CAAB1037.1;--
 DR InterPro: IPR001865; RIBosomal_S2.
 DR PROSITE: PS00962; RIBOSOMAL_S2.1; UNKNOWN_1.
 KW Signal.
 FT SIGNAL.
 FT CHAIN.
 SQ SEQUENCE 1662 AA; 174293 MW; 639EE238600D9246 CRC64;

Query Match 9.3%; Score 288.5; DB 2; Length 1662;
 Best Local Similarity 19.4%; Pred. No. 2,8e-12;
 Matches 133; Conservative 64; Mismatches 144; Indels 345; Gaps 20;

```

OY 19 SSVDARSVAGRSTDMPSGLTKROTQSPPLALYEVLPD-----PLKAP 63
DB 75 SEADALFAGKFTG-----RSATDYGKALRKFIIDPLPGKLSLADGVQRTIPVAVP 128
OY 64 NTVPNNTG-----EDILYYEMELRPSHOIYDLE-PANVGY-----DG----- 103
DB 129 AKWINPOTGAATSD--YEIALVEYKQKLHSDLKMPYTLRGVQLSDGTGLYLYPNS 186
OY 104 -----MSP-----GPTIYPRGTSVY 120
DB 187 SKVAFANADADADAKPVVIMAVDAKGFMTGAKVQARVYMERHMGPLIQAKGTPTRL 246
OY 121 RYVN-----SGENTSP-----NSVHL 136
DB 247 KFNVLPGRAETTVGADGVQYATARNGDIPLDLKSIAHAGLPGDGFTEFTQNRNIHL 306
OY 137 HG-----SESRAPFGMAEDTQPG 156
DB 307 HGGDTPIWISDGTPTQHTPIEEANANPKALVNOGIDPEFLPSFLGASQNVPMDDPG 366
OY 157 E-YKDYYPNRQAARMLYHDAHMSITAEVAGAVYIODPAED-----ALNL 206
DB 367 AGASTYTFPNGQASARMLWYDHITIGVRLNVAGMAAVYTLGDEVDDQLGKTGGLNK 426
OY 207 PSCYGEFDIPVLTAKRY-----NADGTLFSTGEVSSF----- 240
DB 427 VLPPADDTLPVLTDRFVPADVALODARWNTSAMGSESDSWPPHYETVQDPNONGFN 486
OY 241 -----MGDVYVNGOPMMLNVPKRYFR 265
DB 487 SVGRMHMGPFVFPFAMDLPSGEYGDVITVPEAMMDPLVNGAVPTIELDPKYRKM 546
OY 266 FLNAVSRSFALYATSESE-----TRL----- 289
DB 547 VLNASNDREFNISLFLVADEAQRNLNDPLLGATEVKWDAVASATPCAAVGTAVVATDGS 606
OY 290 -----PROVIAADGGL-----EGPVD--- 306
DB 607 YCTPETWPTDNRPGVSPPAQSPFEOIANEGGLPKVAELIAPTPVGYQLDGRITVLN 666
OY 307 --TDTLYISAEEMEVYIDSTPAGOSIDIRNLPGADGLGVEPEFD----- 350
DB 667 VLTGTLGNARADVLVDLSAYAGKTLIYVNDSGAPVPAGDPRNDYFTAVGDQSDAGCA 726
OY 351 -----NTDKVMRFVYDEVLESP 367
DB 727 EDTKPGYGPRTTMMQIKYRAAITTP 752

RESULT 15
O9PME8 PRELIMINARY; PRT; 513 AA.
ID O9PME8;
AC O9PME8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:54:16 ; Search time 25.3 Seconds
(without alignments)
562.851 Million cell updates/sec

Title: US-09-656-640A-2

Perfect score: 3114
Sequence: 1 MISQATGAVALGLAVIGSS.....IQEMASFPYQAQDDDAEE 583

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents-AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PC10S.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3114	100.0	583	4	US-09-401-476-2
2	254.5	84.5	572	4	US-09-401-476-4
3	254.5	8.2	616	1	US-08-749-882A-2
4	254.5	8.2	616	2	US-08-539-134-2
5	253.5	8.1	616	2	US-08-991-531-2
6	253.5	8.1	616	2	US-09-032-315-9
7	253.5	8.1	616	2	US-08-993-318A-9
8	253.5	8.1	616	3	US-09-028-887-2
9	253.5	8.1	616	4	US-09-399-886-9
10	253.5	8.1	616	4	US-09-396-260-9
11	253.5	8.1	616	4	US-09-518-901-2
12	253.5	8.1	616	4	US-09-576-281-9
13	253	8.1	620	1	US-08-940-661A-2
14	253	8.1	620	2	US-09-083-485-2
15	253	8.1	620	2	US-08-939-218A-2
16	252.5	8.1	616	5	PC1-US95-06816-2
17	241	7.7	573	2	US-08-991-531-1
18	241	7.7	573	2	US-09-032-315-10
19	241	7.7	573	3	US-08-993-318A-10
20	241	7.7	573	3	US-09-028-887-1
21	241	7.7	573	4	US-09-399-886-10
22	241	7.7	573	4	US-09-396-260-10
23	241	7.7	573	4	US-09-518-901-1
24	241	7.7	573	4	US-09-576-281-10
25	241	7.7	620	1	US-08-706-037-27
26	241	7.7	620	1	US-09-005-397-27
27	241	7.7	620	5	PC1-US95-06815-2

28	220	7.1	516	3	US-08-689-421-29	Sequence 29, Appl
29	220	7.1	516	4	US-09-389-528-29	Sequence 29, Appl
30	220	7.1	516	4	US-09-181-827A-29	Sequence 29, Appl
31	213	6.8	529	1	US-08-172-318-14	Sequence 14, Appl
32	213	6.8	529	1	US-08-706-037-25	Sequence 25, Appl
33	213	6.8	529	2	US-09-005-397-25	Sequence 25, Appl
34	213	6.8	529	2	US-09-032-315-5	Sequence 5, Appl
35	213	6.8	529	2	US-08-993-318A-5	Sequence 5, Appl
36	213	6.8	529	4	US-09-399-886-5	Sequence 5, Appl
37	213	6.8	529	4	US-09-396-260-5	Sequence 5, Appl
38	213	6.8	529	4	US-09-576-281-5	Sequence 5, Appl
39	202.5	6.5	539	2	US-09-032-315-1	Sequence 1, Appl
40	202.5	6.5	539	2	US-08-993-318A-1	Sequence 1, Appl
41	202.5	6.5	539	3	US-08-689-421-27	Sequence 27, Appl
42	202.5	6.5	539	4	US-09-399-886-1	Sequence 1, Appl
43	202.5	6.5	539	4	US-09-396-260-1	Sequence 1, Appl
44	202.5	6.5	539	4	US-09-389-528-27	Sequence 27, Appl
45	202.5	6.5	539	4	US-09-181-827A-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-09-401-476-2
; Sequence 2, Application US/09401476
; Patent No. 6168936
; GENERAL INFORMATION:
; APPLICANT: Wang, Huaming
; TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes
; FILE REFERENCE: GC584
; CURRENT APPLICATION NUMBER: US/09/401,476
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; TYPE: PR1
; ORGANISM: Stachybotrys chartarum
US-09-401-476-2

Query Match	100.0%	Score 3114;	DB 4;	Length 583;
Best Local Similarity	100.0%	Pred. No. 3.3e-290;		
Matches 583;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MISQATGAVALGLAVIGSSVDARSVAGRSTDMPSGLTRKOTLSPPALYLEVPLPIPL 60		
DB	1	MISQATGAVALGLAVIGSSVDARSVAGRSTDMPSGLTRKOTLSPPALYLEVPLPIPL 60		
QY	61	KAPNTVPNNTGEDIITYEMETIRPESHQIYPLLEPANMVGYSMPGPTIIVRGTESV 120		
DB	61	KAPNTVPNNTGEDIITYEMETIRPESHQIYPLLEPANMVGYSMPGPTIIVRGTESV 120		
QY	121	RFVNSENTSPNSVHLHGFSRAPFDGMAEDTTOPEYKDYYPNROAARMLYHDHMS 180		
DB	121	RFVNSENTSPNSVHLHGFSRAPFDGMAEDTTOPEYKDYYPNROAARMLYHDHMS 180		
QY	181	ITAENAYMOAGYMIQDPAEDALNPSGYGEFDIPLVLTAKRYNADGTLFTNGEVSF 240		
DB	181	ITAENAYMOAGYMIQDPAEDALNPSGYGEFDIPLVLTAKRYNADGTLFTNGEVSF 240		
QY	241	WGDVIOVNOQPMPLNVQKRTKFRFLNAVSKSFLYIATSDSTRLPFOYIADGGL 300		
DB	241	WGDVIOVNOQPMPLNVQKRTKFRFLNAVSKSFLYIATSDSTRLPFOYIADGGL 300		
QY	301	LEGPVDTDLTYISMAERWEVIDESFAGOSTDIRLPGADGIGVPEPDNDKVARFV 360		
DB	301	LEGPVDTDLTYISMAERWEVIDESFAGOSTDIRLPGADGIGVPEPDNDKVARFV 360		
QY	361	DEVLESPTSEVPANLRDVPFPEGGMNDPANPTDDEFTFGGRANGWTINGVTFSDENR 420		
DB	361	DEVLESPTSEVPANLRDVPFPEGGMNDPANPTDDEFTFGGRANGWTINGVTFSDENR 420		

Oy	421	LLRNP/PROTVEIWRLENNNSQMTBPHVHIVDFEVLRSSTARAGPEPEAGLKRQVWLAR	480
	421	LLRNP/PROTVEIWRLENNNSQMTBPHVHIVDFEVLRSSTARAGPEPEAGLKRQVWLAR	480
	421	LLRNP/PROTVEIWRLENNNSQMTBPHVHIVDFEVLRSSTARAGPEPEAGLKRQVWLAR	480
Oy	481	REVVYVVEHVAHPFGGYVWLHCHNLTHEDHQMAAFNPNTVLGDYCYNTTEFIDPHEPLMPR	540
	481	REVVYVVEHVAHPFGGYVWLHCHNLTHEDHQMAAFNPNTVLGDYCYNTTEFIDPHEPLMPR	540
	481	REVVYVVEHVAHPFGGYVWLHCHNLTHEDHQMAAFNPNTVLGDYCYNTTEFIDPHEPLMPR	540
Oy	541	RPFLIGEFENGSGPSELATIDRIQEMASFPNPAQADDDAAEE	583
	541	RPFLIGEFENGSGPSELATIDRIQEMASFPNPAQADDDAAEE	583
	541	RPFLIGEFENGSGPSELATIDRIQEMASFPNPAQADDDAAEE	583

RESULT 2
 US-09-401-476-4
 : Sequence 4, Application US/09401476
 : Patent No. 6168936
 : GENERAL INFORMATION:
 : APPLICANT: Wang, Huaming
 : TITLE OF INVENTION: No. 6168936el Phenol oxidizing Enzymes
 : FILE REFERENCE: GC584
 : CURRENT APPLICATION NUMBER: US/09/401,476
 : CURRENT FILING DATE: 1999-09-22
 : NUMBER OF SEQ ID NOS: 8
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 4
 : LENGTH: 572
 : TYPE: PR1
 : ORGANISM: Stachybotrys charatum
 : US-09-401-476-4

Query Match	64.5%;	Score 2010;	DB 4;	Length 572;
Best Local Similarity	65.4%;	Pred. No. 2.9e-184;		
Matches 383;	Conservative 71;	Mismatches 110;	Indels 22;	Gaps 10;

[illegible]

```
QY      533 PMEPLMRPRPFLLEGFEENGSGDFSELATTDRIQEMASFNPYAQADD 578
      ||| ||: ||: ||||: || || ||: |||| ||: ||| ||:
Db      527 PMEELQARPYELGEEFQAQSGGSVAQVTERIQTMAEYRPPYAAADE 5722
```

RESULT 3
 US-08-749-882A-2
 Sequence 2, Application US/08749882A
 Patent No. 5750388
 GENERAL INFORMATION:
 APPLICANT: Berka, Randy
 APPLICANT: Thompson, Sheryl
 APPLICANT: Xu, Feng
 TITLE OF INVENTION: Purified Scytalidium Laccases
 TITLE OF INVENTION: And Nucleic Acids Encoding Same
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 5750388disk Of No. 5750388eth America, Inc.
 STREET: 405 Lexington Avenue - 64th Fl.
 CITY: New York
 STATE: NY
 COUNTRY: USA
 Zip: 10174
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/749,882A
 FILING DATE: 15-NOV-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4166,020-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-878-9652
 TELEFAX: 212-878-9655
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 616 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE: Scytalidium thermophilum
 ORGANISM: Scytalidium thermophilum
 US-08-749-882A-2

Query Match	8.2%;	Score 254.5;	DB 1;	Length 616;
Best Local Similarity	25.2%;	Pred. No. 9.8e-16;		
Matches 137;	Conservative 71;	Mismatches 171;	Indels 165;	Gaps 33;

```

QY 69 PNTGDILEYEMERPRPSHOIYPDLGEPANNVGYGMSR-----GPTIIVRGTE 117
Db 74 PNTG-----VYRKTFPI-----TEVDNRRGPGVIKERLMLINDKLLOGPTYPANNG-D 121
QY 118 SVAFVNGSEITSPNSVYHLHGSFRAP-----FDGMAEDTTOGGEYKDYYPNNQOARML 172
Db 122 TIEVTYNNHLRINTGTSIIHMGHLOKGTNYHDGANGVYECIEIPRGSRYVSFRARQVTS- 180
QY 173 WYHDHAMSITAENNAVMQAGVYMIODPAEDALNIPSGYGFDPPL-VLTAKRY--NADG 228
Db 181 WYHSH-PSAQGN---GVSGAIQIINGSA-----SLP-----YDIDIGVLPIDMWYKNSADQ 227
QY 229 TLESTNEVSSFWGDVLYQVNGQ-----PMDMLNVQ-RTYRFFFLAANYSRSPALY 278
Db 228 LVETTLAKGNAPESDYNVLINGTAKHPTTGGEYVAIVLTLDPKRRRLRLIMNSVENHFQVS 287
QY 279 LATSEDETRLPFOVIAADGGLLEGPAVD-----TTLVYLSMAERWEVVDIES----- 325

```

Db 288 LA-----KHTMTVIAD-----MVPVAMTVDLSLFAVAGORYDVTIDASQAVGNWYFNI 336
 QY 326 TFAGOS-----IDIRNLPGA-DGLGVEPEFNDTKVMRFVNDVELESPTSEVPA 374
 Db 337 TFGGQOKGCFSHNPAPAIRFYEGAPDALPTDGAAPKD-----HQCDDTLDLS--PV 387
 QY 375 MLRDVP-----FPEGGNMPPANPTDDE-----TFTEGRANGQWTINGTFSQVDE-----418
 Db 388 VQKNVPYDGFVKERGNLPTLHYDQAAPRHVF-----WKING--SADVDMDRVPYLE 439
 QY 419 ---NRLRNVP--RDYVEI-----WRLNNSNG---WTHPVHILVDPRVLSRS---459
 Db 440 YVMNNDLSIPKNNIYRVGDNEMTYWLVNDEGRSLSPHPMHLGHDFEVLGRSPDV 499
 QY 460 -----TARGVEPEYEAAGLKDYVWVLARREVVYVEAHYAPPGVYMLHCH 502
 Db 500 SPDETFEVDPAVDLRLRGHNPVR-----RDVTMLPARGWLL-ARTDNPGMALHCH 554
 QY 503 NLIH 506
 Db 555 IAWH 558

RESULT 4

US-08-539-134-2
 ; Sequence 2, Application US/08539134
 ; Patent No. 5843745
 ; GENERAL INFORMATION:
 ; APPLICANT: Beika, Randy
 ; APPLICANT: Thompson, Sheryl
 ; APPLICANT: Xu, Feng
 ; TITLE OF INVENTION: Purified Scytalidium laccases
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 58437450 No. 5843745disk Of No. 5843745th America, Inc.
 ; STREET: 405 Lexington Avenue - 64th Fl.
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/539,134
 ; FILING DATE: 04-OCT-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lambiris, Elias J
 ; REGISTRATION NUMBER: 33,728
 ; REFERENCE/DOCKET NUMBER: 4186.010-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-878-9652
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 616 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Scytalidium thermophilum
 ; US-08-539-134-2

Query Match 8.28; Score 254.5; DB 2; Length 616;
 Best Local Similarity 25.28; Freq. No. 9.0e-16;
 Matches 137; Conservative 71; Mismatches 171; Indels 165; Gaps 33;

QY 69 PNTGEDILYYEMELRPSHQIYPLDEBANNVGYDGMSF-----GPTIIVPGTE 117
 Db 74 PNTG-----VYRRTFPDI---TEVDNRRGPDGVIKREKMLINDKLLGPTVANNNG-D 121
 QY 118 SVAFVSGENTSPNSVHLHSGSRAP-----FDGMAEDTTPGCEYKDYIYPPNQAARML 172
 Db 122 TIEVTVNNHLTNGTSTIHMGHKGCTNYHIDGANGVTECPRPBGSHVYSFRARQYCTS- 180
 QY 173 MYDHAASTIAENAYMGOAGYVMJODPAEDALNLPSTYGERDIPL-VLTAKRY---NADG 228
 Db 181 WYHSH-FSAOYGN---GVSGAIQINGPA---SLP-----YDIDGLVPLQDMWYTKSADQ 227
 QY 229 TLFSTNGEVSFWGDVIOVNGQ-----PMPLNVOP-RKYRFFELNAAYSRSEFALY 278
 Db 228 LVITETLAKGNAPFSDNVNLINGTAKHPTTGEGEYAIYKLTDPKRRRLLLINNSVENHFOYS 287
 QY 279 LATSDESETRLPFOVYIADGGLLEGVD---TDTLYISMAERWEVYIDFS-----325
 Db 288 LA-----KHTMTVIAD-----MVPVAMTVDLSLFAVAGORYDVTIDASQAVGNWYFNI 336
 QY 326 TFAGOS-----IDIRNLPGA-DGLGVEPEFNDTKVMRFVNDVELESPTSEVPA 374
 Db 337 TFGGQOKGCFSHNPAPAIRFYEGAPDALPTDGAAPKD-----HQCDDTLDLS--PV 387
 QY 375 MLRDVP-----FPEGGNMPPANPTDDE-----TFTEGRANGQWTINGTFSQVDE-----418
 Db 388 VQKNVPYDGFVKERGNLPTLHYDQAAPRHVF-----WKING--SADVDMDRVPYLE 439
 QY 419 ---NRLRNVP--RDYVEI-----WRLNNSNG---WTHPVHILVDPRVLSRS---459
 Db 440 YVMNNDLSIPKNNIYRVGDNEMTYWLVNDEGRSLSPHPMHLGHDFEVLGRSPDV 499
 QY 460 -----TARGVEPEYEAAGLKDYVWVLARREVVYVEAHYAPPGVYMLHCH 502
 Db 500 SPDETFEVDPAVDLRLRGHNPVR-----RDVTMLPARGWLL-ARTDNPGMALHCH 554
 QY 503 NLIH 506
 Db 555 IAWH 558

RESULT 5

US-08-991-531-2
 ; Sequence 2, Application US/08991531
 ; Patent No. 5925554
 ; GENERAL INFORMATION:
 ; APPLICANT: Pedersen, Anders Hjelholt
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Schneider, Palte
 ; APPLICANT: Rasmussen, Grethe
 ; APPLICANT: Cherry, Joel
 ; TITLE OF INVENTION: Myceliophthora And Scytalidium Laccase
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 59255540 No. 5925554disk of No. 5925554th America, Inc.
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/991,531
 ; FILING DATE: 16-DEC-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rozek, Carol
 ; REGISTRATION NUMBER: 36,993

REFERENCE/DOCKET NUMBER: 5125.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-531-2

Query Match 8.1%; Score 253.5; DB 2; Length 616;
Best Local Similarity 25.2%; Pred. No. 1.2e-15;
Matches 137; Conservative 70; Mismatches 172; Indels 165; Gaps 33;

QY 69 PNTGEDILYEMERPESHQIYDLEPANWGYDMSPP-----GPTIIVRGTE 117
DB 74 PNTG-----VYRRYTPDI---TEVDNRPDGVYKEKMLINDKLGPVFANWG-D 121

QY 118 SVYRFVNGENTSPNSVHLHSGFSRAP-----FDGNAEDTQGEYKDYYPNRQARML 172
DB 122 TIEVTVNNHLRTGTSIHWHGLHOKGTNYHDGANGVTECPITPGGSRVYSFRARQYGS- 180

QY 173 WYDHAMSITAENAVYGOAGVYMIODPAEDALNPSGYGFEDIP-L-VLTAARY---NADG 228
DB 181 WYHSH-FSAQYGN---GVSGAIQINGPA---SLP-----YDIDGLVPLXDMVYKSADQ 227

QY 229 TLESTNGEVSSEFGDYIQUVNGQ-----PWPMLNVP-RKYRFRFLAAVSRSFALY 278
DB 228 LVETELXKGNAPPSDVNLINGTAKHPTTGEYAIYKLPDKRHRRLINMSVENHFOVS 287

QY 279 LATSSESETRLPFOVIAADGLEGVD---TDLYTISMAERWEVIDFS----- 325
DB 288 LA-----KHTMTVIAAD---MVPVNAMTVVSLFMAVGQRYDVTIDASQAVGNWENI 374

QY 326 TFGAGS-----IDIRNLPGA-DGLGVEPEFNTDKVMRFVVDVLESPDTSEVPA 374
DB 337 TFGGQKCGFSHPADPAIFRYEGAPDALPTDGAAPKD-----HQCIDLTLDS--PV 387

QY 375 NLBDVP---FPEGGMWDANPTDDE---TTFGRANQMTINGVTSQVE----- 418
DB 388 YOKNVPDGVFKRPGNTLPVTLHVDAQAAPHVT-----WKLING-SAADVWMDRPVLE 429

QY 419 ---NRLLRNP--RDYVEI-----WRLSNSNG---WTHVPHILVDFRVLRS--- 459
DB 440 YVANNLSSIPYKNNIVRVGQVNEWTYWLVEPDEGRSLPHPMHLHGHDFFVLGRSPDV 499

QY 460 -----TARGVEPYEAGLKDYVWLARREVYVEAHYAPFPGVYMLHCH 502
DB 500 SPDSETRFVFPDPAVDLPRLRGHNPVR---RDVTMLPARQWMLL-AFRDNPQAWLFFCH 554

QY 503 NLH 506
DB 555 IAXH 558

RESULT 6
US-09-032-315-9
Sequence 9, Application US/09032315
Patent No. 5985818
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59858180 No. 5985818disk of No. 5985818ch America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,315
FILING DATE: 27-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 5200.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-315-9

Query Match 8.1%; Score 253.5; DB 2; Length 616;
Best Local Similarity 25.2%; Pred. No. 1.2e-15;
Matches 137; Conservative 70; Mismatches 172; Indels 165; Gaps 33;

QY 69 PNTGEDILYEMERPESHQIYDLEPANWGYDMSPP-----GPTIIVRGTE 117
DB 74 PNTG-----VYRRYTPDI---TEVDNRPDGVYKEKMLINDKLGPVFANWG-D 121

QY 118 SVYRFVNGENTSPNSVHLHSGFSRAP-----FDGNAEDTQGEYKDYYPNRQARML 172
DB 122 TIEVTVNNHLRTGTSIHWHGLHOKGTNYHDGANGVTECPITPGGSRVYSFRARQYGS- 180

QY 173 WYDHAMSITAENAVYGOAGVYMIODPAEDALNPSGYGFEDIP-L-VLTAARY---NADG 228
DB 181 WYHSH-FSAQYGN---GVSGAIQINGPA---SLP-----YDIDGLVPLXDMVYKSADQ 227

QY 229 TLESTNGEVSSEFGDYIQUVNGQ-----PWPMLNVP-RKYRFRFLAAVSRSFALY 278
DB 228 LVETELXKGNAPPSDVNLINGTAKHPTTGEYAIYKLPDKRHRRLINMSVENHFOVS 287

QY 279 LATSSESETRLPFOVIAADGLEGVD---TDLYTISMAERWEVIDFS----- 325
DB 288 LA-----KHTMTVIAAD---MVPVNAMTVVSLFMAVGQRYDVTIDASQAVGNWENI 374

QY 326 TFGAGS-----IDIRNLPGA-DGLGVEPEFNTDKVMRFVVDVLESPDTSEVPA 374
DB 337 TFGGQKCGFSHPADPAIFRYEGAPDALPTDGAAPKD-----HQCIDLTLDS--PV 387

QY 375 NLBDVP---FPEGGMWDANPTDDE---TTFGRANQMTINGVTSQVE----- 418
DB 388 YOKNVPDGVFKRPGNTLPVTLHVDAQAAPHVT-----WKLING-SAADVWMDRPVLE 429

QY 419 ---NRLLRNP--RDYVEI-----WRLSNSNG---WTHVPHILVDFRVLRS--- 459
DB 440 YVANNLSSIPYKNNIVRVGQVNEWTYWLVEPDEGRSLPHPMHLHGHDFFVLGRSPDV 499

QY 460 -----TARGVEPYEAGLKDYVWLARREVYVEAHYAPFPGVYMLHCH 502
DB 500 SPDSETRFVFPDPAVDLPRLRGHNPVR---RDVTMLPARQWMLL-AFRDNPQAWLFFCH 554

QY 503 NLH 506
DB 555 IAXH 558

APPLICATION NUMBER: US/09/032,315
FILING DATE: 27-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 5200.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ. ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-396-260-9

Query Match 8.1%; Score 253.5; DB 4; Length 616;
Best Local Similarity 25.2%; Pred. No. 1,2e-15;
Matches 137; Conservative 70; Mismatches 172; Indels 165; Gaps 33;

QY 69 PNTGEDILYEMETIRPESHOIYDLEPANNVGYDGMSPP-----GPTIIVPGTE 117
DB 74 PNTG-----VYRRYTFDI---TEVDNRPGRPDGVYKEKMLINDKLGPTEFANMG-D 121
QY 118 SVYRFVNGENTSPNSVYHLHGSFSRAP-----FDGNAEDTTPGEEKDYYPNROAARML 172
DB 122 TIEVTYVNNHRLRTGTSIHMHGKHOKGTNYHDGANGVTECPYIPGGSRYVSFRARQYCTS- 180
QY 173 WYDHAMSTIAENAYMOAGVYMIODPAEDALNLPSSYGFEDIPL-VLTAKRY---NADG 228
DB 181 WYHSH-FSAQYGN---GVSGAIOINGPA-----SLP-----YDIDGLVPLXDMYKRSADQ 227
QY 229 TLFSTNGEVSSFGVDYIQVNGQ-----PWPMLNVQP-RKYRFRFLNAVSRSFALY 278
DB 228 LVETETLKGNAPPSDNLVINGTAKHPTTGEGEYAIYVLTPTDKRHRLRLIMMSVENHFOVS 287
QY 279 LATESESETRLPFOVYIADGGLLEGPDV---TDTLYISMAERWEVYIDFS----- 325
DB 288 LA-----KHTMTVIAAD---MVPYNAFTVSLFMAVQORYDVTIDASQAVGNWFNI 336
QY 326 TFGAGQS-----IDIRNLPGA-DGLGVEPEFNDTKVMRFVYVDEVLESPTSEVPA 374
DB 337 TFGGQKCGFSHPNAPARAIRYEGAPDALPTDGAAPKD-----HQCIDTIDL-S--PV 387
QY 375 NLADVP---FPEGKMDPANPTDDE---TTFGRANGQMTINGVTSQVDE----- 418
DB 388 VQKNVPYDGFVKBERGNTLPTLVHVDQAARHVFVFT-----WKING-SAADVDMREVLE 439
QY 419 ---NRLLRNVP---RDYVEI-----WRLNNSNG---WTHPVHILVDFRVLRS- 459
DB 440 YVANNNDLSSIPVKNNTIVRYDGVNEMTYWLVENDEBGRSLSPHPRHILGHDFVILGRSPDV 499
QY 460 -----TANGVERTEAAGLKDYVWMLARREYVYVEAHYAFPGVYMLHCH 502
DB 500 SPDSETRFVDPVAVDLPRLGHNVPV---RDVTMLPARQWMLL-ARTDNPQAWLFRCH 554
QY 503 NLIN 506
DB 555 IAXH 558

RESULT 11

US-09-518-901-2
Sequence No. Application US/09518901
Patent No. 6218170
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6218170d No. 6218170disk of No. 6218170th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/518,901
FILING DATE: 06-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/028,887
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 5201.200-US
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-518-901-2

Query Match 8.1%; Score 253.5; DB 4; Length 616;
Best Local Similarity 25.2%; Pred. No. 1,2e-15;
Matches 137; Conservative 70; Mismatches 172; Indels 165; Gaps 33;

QY 69 PNTGEDILYEMETIRPESHOIYDLEPANNVGYDGMSPP-----GPTIIVPGTE 117
DB 74 PNTG-----VYRRYTFDI---TEVDNRPGRPDGVYKEKMLINDKLGPTEFANMG-D 121
QY 118 SVYRFVNGENTSPNSVYHLHGSFSRAP-----FDGNAEDTTPGEEKDYYPNROAARML 172
DB 122 TIEVTYVNNHRLRTGTSIHMHGKHOKGTNYHDGANGVTECPYIPGGSRYVSFRARQYCTS- 180
QY 173 WYDHAMSTIAENAYMOAGVYMIODPAEDALNLPSSYGFEDIPL-VLTAKRY---NADG 228
DB 181 WYHSH-FSAQYGN---GVSGAIOINGPA-----SLP-----YDIDGLVPLXDMYKRSADQ 227
QY 229 TLFSTNGEVSSFGVDYIQVNGQ-----PWPMLNVQP-RKYRFRFLNAVSRSFALY 278
DB 228 LVETETLKGNAPPSDNLVINGTAKHPTTGEGEYAIYVLTPTDKRHRLRLIMMSVENHFOVS 287
QY 279 LATESESETRLPFOVYIADGGLLEGPDV---TDTLYISMAERWEVYIDFS----- 325
DB 288 LA-----KHTMTVIAAD---MVPYNAFTVSLFMAVQORYDVTIDASQAVGNWFNI 336
QY 326 TFGAGQS-----IDIRNLPGA-DGLGVEPEFNDTKVMRFVYVDEVLESPTSEVPA 374
DB 337 TFGGQKCGFSHPNAPARAIRYEGAPDALPTDGAAPKD-----HQCIDTIDL-S--PV 387
QY 375 NLADVP---FPEGKMDPANPTDDE---TTFGRANGQMTINGVTSQVDE----- 418
DB 388 VQKNVPYDGFVKBERGNTLPTLVHVDQAARHVFVFT-----WKING-SAADVDMREVLE 439
QY 419 ---NRLLRNVP---RDYVEI-----WRLNNSNG---WTHPVHILVDFRVLRS- 459
DB 440 YVANNNDLSSIPVKNNTIVRYDGVNEMTYWLVENDEBGRSLSPHPRHILGHDFVILGRSPDV 499
QY 460 -----TANGVERTEAAGLKDYVWMLARREYVYVEAHYAFPGVYMLHCH 502
DB 500 SPDSETRFVDPVAVDLPRLGHNVPV---RDVTMLPARQWMLL-ARTDNPQAWLFRCH 554

QY 503 NLH 506
DB 555 IAXH 558

RESULT 12
US-09-576-281-9
Sequence 9, Application US/09576281
Patent No. 6277611
GENERAL INFORMATION:
APPLICANT: Pedersen, Anders
APPLICANT: Svendsen, Allan
APPLICANT: Schneider, Palle
APPLICANT: Rasmussen, Grethe
APPLICANT: Cherry, Joel
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 62776110 No. 6277611disk of No. 6277611ch America
STREET: 405 Lexington Avenue
CITY: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/576,281
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,318
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Greg9, Valeta A.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5032.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-576-281-9

Query Match 8.1%; Score 253.5; DB 4; Length 616;
Best Local Similarity 25.2%; Pted. No. 1.2e-15;
Matches 137; Conservative 70; Mismatches 172; Indels 165; Gaps 33;

QY 69 PNTGEDIYEMIEIRPFSSHOYIPDLBPANVGYDGMSPP-----GPTIIVPGTE 117
DB 74 PNTG-----VYKRTPTI---TEVDNRPFGDGYIKKMLINDKLGPTYPANMG-D 121
QY 118 SVRVNNGENTSPNSVHLHGSFSRAD-----FDGMAEDTTPGEXKDYVYVPRQAARM 142
DB 122 TIEVTANNHRTNGTSTHWHGLHQKGTNYHDGANGVTECPIPPGGSFVYSFRAROGTS- 180
QY 173 WYDHDAMSTIAENAYMGQAGVYMIODPAEDALMLPSGYGEFDIPL-VLTAKRY---NADG 228
DB 181 WYSH-FSAOYGN--GVSAGIQINGPA---SLP-----YDIDGLVLPXDMVYKSAQ 227
QY 229 TLFSTNGEVSFPGDVIVQNGO-----PWPMLNVOP-RKYRPFRLNAVSRSPALY 278
DB 228 LVLETLLAKGNAPSDVNLINGTAKHPTTGGEYALVYKLTDPKRRRLRLIMSVENHFOVS 287

QY 279 LATSEDESTRLPFOVIADGGLGEPVD---TDTLYISAEWEVVIDS----- 325
DB 288 LA-----KHTMTVIAAD-----MVPNAMTVDSLEFMAVGORVDYIDASAGVYWNFI 336
QY 326 TFGAGS-----IDIRNLPGA-DGLGVEPEPNTDKVAREFVYDEVLESDTSEVA 374
DB 337 TFGGQCKGCFSHNPAPAAIFREYEGAPDALPTDGAAPKD-----HQLDITLDS--EV 387
QY 375 NLRDVP-----FPEGWMDPANPTDDE-----PFTFGANGOWTINCFTSDVE----- 418
DB 388 VOKNVFVDGFEKPGNTLPYTLHAVDOAAPHYFT-----WKING-SAADDVMDRPYLE 439
QY 419 ---NRLRNVP--RDVEI-----WLENNNSG---WTHPVHILVDFYLSRS--- 459
DB 440 YVANNLSSISPVKNINIVRVGVNWEWTYLVLENDPEGRLSLPHRMHLHGHDFFVLGRSPDV 499
QY 460 -----TARGVEPYEAGLKQVYVLAAREVYVYEAHPFQVYVYLLHCH 502
DB 500 SPDSETRFVFPADVLPRLGNHVPV---RDVTMLPARGWILL-AFRDNPQAWLFFCH 554
QY 503 NLH 506
DB 555 IAXH 558

RESULT 13
US-08-940-661A-2
Sequence 2, Application US/08940661A
Patent No. 5795760
GENERAL INFORMATION:
APPLICANT: BERKA, Randy Michael
APPLICANT: BROWN, Stephen H.
APPLICANT: XU, Feng
APPLICANT: SCHNEIDER, Palle
APPLICANT: OXENB LL, Karen M.
TITLE OF INVENTION: PURIFIED MYCELIOPHTHORA LACCASES AND NUCLEIC
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 57957600 No. 5795760disk of No. 5795760th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940,661A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/441,146
FILING DATE: 15-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 4184.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 620 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-940-661A-2

Query Match	8.1%	Score 253;	DB 1;	Length 620;
Best Local Similarity	25.28;	Pred. No. 1.4e-15;		
Matches 153; Conservative	56;	Mismatches 229;	Indels 168;	Gaps 31

QY	66	VPNNTGCDILLYEMELRPESHOLYDPLEBANMVGCGMSP-----GPTLIVPR	114
Db	72	VDSPTG-----VVRPTLTL---TEVDNMTGDPGVYKEMVLYNNSIGITPLADW	120
QY	115	GTESVVRFVRENGENTSPNSVHLHGSFSRAP--FDGMAEDTTP-----GEKRDYYPNRQA	168
Db	121	GDYIQQVYINLLE-TNGTSHMHGHLQKGTNLHDGANGITECPJPRPGKRYVRFKAQY	179
QY	169	ARMIMYHDHAMSITAENAYMQAQGVYMIODPAEDALMLPSGYGEFDI-----PLYL	219
Db	180	GTS-WYHSH-FSAQYGN--GVWGAIDINGNA-SLEPYDLDGAFPSIDYYSASADELYE	232
QY	220	TAKRYNA-----DGTLES-----TNGEVSFSGWDVIOYNGQPMWPLANTQPKRYRFLNAA	270
Db	233	LTKNSGAFPSQNVLFNCGIAKHPTGE-----GEYANTTLRP-----GRRHRLRLNMS	280
QY	271	VSRSPALYLATSEDESTRLPFOVITADGGLLEGVD--TDLYISMAERMEVYVDFS-	325
Db	281	VENHFQVSLVNH-----TWTIIAAD--MVPYNAWMTDSLFLTGQGRDYVIEASRT	329
QY	326	-----TFAGOSI-----DIRNLPGADGLVEEPFDOTDKVMFVYDEVLESP	367
Db	330	PGNTWENVTFFGGGLCGGSRNPYPALAFHAGA-----PGRPPTDGAKPAVDHNCILDP	383
QY	368	DTSEVPANLRVFPPEPEGNMDPANPTDETFTEFRANGOWTINVTFPS-----DYENRL	421
Db	384	NLKPYVA--RDVPL-SGPARKPDNTLDVTLDTGTPLFVWKVNSAINTIDMGRVYDYL	440
QY	422	LRRNP-----RDYELIRLENNSG---WHPVPHILVDPRVLSRSTANG--	463
Db	441	TONTSPFGYNIIVEVNGADOMSYMLIENDPGAPTLPHPMHLHGDFYVLCGRSPDESAS	500
QY	464	-----VEPYEAGLKDVVWMLARREV-----VEAAYAPFPGVYMLHCNLIHEBDM	511
Db	501	NERHVFDDARAGLLSGANPYRRDYTMPLPAFGWVYLLAFRADNPGAWLFRCHITAMH-----	555
QY	512	MAAFNVTVLGDYGVNYTEFIDMPELRMPREPLFLGEEFENGSGDESELAITDRIQEMASFN	571
Db	556	-----VSGELGVYLLERAAD-----LRGAVSADADADDLDRLCAWMRYMYPT-N	597
QY	572	PYAQAD	577
Db	598	PYPKSD	603

RESULT 14
US-09-083-485-2
Sequence 2, Application US/09083485
Patent No. 5948121
GENERAL INFORMATION:
APPLICANT: Dorrit Aaslyng
APPLICANT: Sorensen, Niels H.
TITLE OF INVENTION: Lacases with Improved Dyeing
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 59481210 NO. 5948121disk OF NO. 5948121th America, Inc
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,485

1 FILING DATE: 20-May-1998
2
3 CLASSIFICATION: 435
4
5 ATTORNEY/AGENT INFORMATION:
6 NAME: Gregg, Valera
7 REGISTRATION NUMBER: 35,127
8 REFERENCE/DOCKET NUMBER: 4639.204-US
9
10 TELECOMMUNICATION INFORMATION:
11
12 TELEPHONE: 212-867-0123
13
14 TELEFAX: 212-878-9655
15
16 INFORMATION FOR SEQ ID NO: 2:
17
18 SEQUENCE CHARACTERISTICS:
19
20 LENGTH: 620 amino acids
21
22 TYPE: amino acid
23
24 TOPOLOGY: linear
25
26 MOLECULE TYPE: protein
27
28 US-09-083-485-2

Query March	8.1%;	Score 253;	DB 2;	Length 620;
Best Local Similarity	25.2%;	Pred. No. 1.4e-15;		
Matches 153;	Conservative 56;	Mismatches 229;	Indels 168;	Gaps 31.

QY	66	VPNNTGSDILLEYEMELRPFESHQIYLPDLERPNANVYGQMSD-----	-----GPTLIVPR	114
Db	72	VDSDPDTG-----VVRPYTLTL-----TEVDNMTGPRGGVYKAKMYLVNNSIIGPTIADW	1200	
QY	115	GTESEVREPVNSCENTSPNSVHLHGSFSRAP--FDGMAEDTTP-----GEYKDYUYPNRQA	168	
Db	121	GDTIQVWYINLLE--TNGTSIHMHGLHOGTMLHDGANGITIECPRLPRGGRKRYVFRKAOX	179	
QY	169	ARMIMYHDHANSITAEANVYMOGAQVYMIOPADEBALMLPSYGEFDI-----PLYL	219	
Db	180	GTS--WYHSH--FSAOYGN--GVVAIDINGA--SLPYDIDLGFPISDIYYSADLVE	232	
QY	220	TAKRYNA-----DGLTFS-----TNGEYSSFEWGDVIOYNGOPWPMYLNQPRKYRFRFLNAA	270	
Db	223	LTKNSGAFFSDNVLFNGAKHPETGE-----GEYANVTLP-----GRRHRLRLNTS	280	
QY	271	VSRFALYLTATSESETRPLRPQVLIADGGCLLEGVD--TDLTLYISAEKMEVYIDS--	325	
Db	281	VENHFOVSLVNH-----TWTIILAAD--MRYVANMTVDSLEFLGAGORADYVIEASRT	329	
QY	326	-----TFAGQSI-----DIRNLPGAGLGLEVEPFMDTKVMRFYVDEVLESP	367	
Db	330	PGNTWENVTFCGGGLCCGSRNPYPAALFHTAGA-----PGRPETDGAKARVHDHNCIDLPR	383	
QY	368	DTSEVPANLRDQVPEPEGNMDPANPTDEDTFTFEGRANGOWTINCVTPS-----DYENRL	421	
Db	384	NLKRPYVA--RQVPL--SGRAKRPDNTLDYTLDTGTPLFVKKVYNSAIALINDMGRVYDYL	440	
QY	422	LRRNP-----RDVIEIRLENNNSNG--WTHPVHILHVDREVLSRSTARG--	463	
Db	441	TQNTSFPFGYNIIVEYNGADQMSYMLIENDPCAPETTLRHPMLHLHDFVYLGSRSPDES	500	
QY	464	-----VEPYEAGLKDQVWMLARRVV-----VYEANVYAPRFQVYMLHCHNLIHEDHM	511	
Db	501	NERHVFDPARAGGLSGCANPYRQRYTMLPAFGWVYLLFRADNFCAMLEPHGHIAH-----	555	
QY	512	MAEFNVTLGDYGYNYTEFIDPMEPLMRPRPFLIGFEENGSGPSELAITDRIOEMASFN	571	
Db	556	-----VSGGLGVVYLERADD-----LHGAVSDADADDLDRLCADWRRYWTPT--N	597	
QY	572	RYAOAD	577	
Db	598	PYPKSD	603	

RESULT 15
US-08-939-218A-2
; Sequence 2, Application US/08939218A
; Patent No. 5981243
; GENERAL INFORMATION:
; APPLICANT: BEKA, Randy Michael

0000

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:54:16 ; Search time 58.78 Seconds
(without alignments)
1101.668 Million cell updates/sec

Title: US-09-656-640A-2

Perfect score: 3114

Sequence: 1 MISQALGAVNALGLAVIGSS.....IQEMASFPYQAQDDAAEE 583

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_032802:*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3114	100.0	583	22	AAB81505	Stachybotrys chart
2	3114	100.0	583	22	AAB20097	Stachybotrys chart
3	3090	99.2	578	22	AAB81506	Bilirubin oxidase.
4	2010	64.5	572	14	AAR40843	Bilirubin oxidase.
5	1829	58.7	602	21	AAV69204	Amino acid sequenc
6	1762	56.6	627	21	AAV95538	Bipolaris spicifer
7	1762	56.6	627	21	AAV96762	Bipolaris spicifer
8	1723.5	55.3	627	21	AAV95539	Curvularia pallasc
9	1723.5	55.3	627	21	AAV96763	Curvularia pallasc
10	1716.5	55.1	594	20	AAV45222	Stachybotrys chart
11	1716.5	55.1	594	20	AAV39992	Stachybotrys pheno

12	1716.5	55.1	594	21	AAV95537	Stachybotrys chart
13	1716.5	55.1	594	21	AAV96761	Stachybotrys chart
14	605	19.4	474	22	AAU04831	Micromonospora eve
15	504	16.2	511	22	AAV90799	C glutamicum prote
16	498.5	16.0	497	22	AAV79105	Corynebacterium gl
17	454	14.6	570	21	AAV44482	Arabidopsis thalia
18	454	14.6	574	21	AAV44481	Arabidopsis thalia
19	454	14.6	581	21	AAV44480	Arabidopsis thalia
20	453	14.5	477	22	AAV83078	S. epidermidis ope
21	355.5	11.4	410	22	AAV83082	S. epidermidis ope
22	342	11.0	113	21	AAV95540	Amerosporium atrum
23	342	11.0	113	21	AAV96764	Amerosporium atrum
24	313	10.1	348	22	AAV82886	S. epidermidis ope
25	254.5	8.2	616	18	AAV18060	Scytalidium thermo
26	253.5	8.1	616	19	AAV76314	Scytalidium thermo
27	253.5	8.1	616	19	AAV59915	Scytalidium thermo
28	253.5	8.1	616	19	AAV62502	Scytalidium thermo
29	253.5	8.1	616	19	AAV51783	Scytalidium thermo
30	253.5	8.1	620	21	AAV91086	Scytalidium thermo
31	253	8.1	620	18	AAV19855	Myceliophthora the
32	243	7.8	573	19	AAV79079	Myceliophthora the
33	241	7.7	573	19	AAV76315	Myceliophthora the
34	241	7.7	573	19	AAV59913	Myceliophthora the
35	241	7.7	573	19	AAV62503	Myceliophthora the
36	241	7.7	573	19	AAV51782	Myceliophthora the
37	241	7.7	573	19	AAV63124	Myceliophthora the
38	241	7.7	573	21	AAV91085	Myceliophthora the
39	241	7.7	620	17	AAV88500	Myceliophthora the
40	241	7.7	620	18	AAV16302	Myceliophthora the
41	241	7.7	620	21	AAV50729	M. thermophila lac
42	226.5	7.3	573	19	AAV76316	Myceliophthora the
43	226.5	7.3	573	19	AAV79077	Myceliophthora the
44	225	7.2	573	19	AAV76317	Myceliophthora the
45	225	7.2	573	19	AAV79078	Myceliophthora the

ALIGNMENTS

RESULT 1	
AA81505	18-JUN-2001 (first entry)
AA81505	Stachybotrys chartarum phenol oxidase B enzyme.
AA81505	Stachybotrys phenol oxidase B; spob; phenol oxidizing enzyme;
AA81505	detergent; paper production; pulp production; textile; food industry;
AA81505	bleaching.
AA81505	Stachybotrys chartarum.
AA81505	WO200121748-A1.
AA81505	29-MAR-2001.
AA81505	06-SEP-2000; 2000WO-EP08840.
AA81505	22-SEP-1999; 99EP-0203120.
AA81505	(UNITL) UNILEVER NV.
AA81505	(UNITL) UNILEVER PLC.
AA81505	(HIND-) HINDUSTAN LEVER LTD.
AA81505	Convents D, Doornik M, De Vries CH, Wang H;
AA81505	WPI, 2001-273462/28.
AA81505	N-PSDB; AAF82586.
AA81505	New detergent compositions comprising a phenol oxidizing enzyme useful

PT in detergent or cleaning compositions, fiber treatment, processing,
PT finishing or production, paper and pulp production, or in starch
PT processing applications

XX Claim 1; Fig 2; 46pp; English.

CC The present sequence is a Stachybotrys chartarum phenol oxidizing
CC enzyme. The invention relates to detergent compositions comprising one or
CC more surfactants and a phenol oxidizing enzyme having at least 68%
CC identity to the Stachybotrys chartarum phenol oxidizing enzyme. Phenol
CC oxidizing enzymes may be used in the detergent, paper, pulp, textile and
CC food industries. They are used for preventing the transfer of dyes in
CC solution from one textile to another during detergent washing, or in
CC modifying the colour associated with dyes and coloured compounds having
CC different chemical structures, such as in pulp and paper bleaching,
CC bleaching the colour of stains on fabric and in detergent and textile
CC applications.

XX Sequence 583 AA;

Query Match 100.0%; Score 3114; DB 22; Length 583;

Best Local Similarity 100.0%; Pred. No. 8.9e-275;

Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISOAIGAVAGLAVIGSSVDARSVAGRSTDMPSGLTKRQTLSPPLALYEVLPIPL 60
DB 1 misqalgaavalglavlgssvdarsvagrstdmgsqtkrqtlspplalylevplippl 60

QY 61 KAPNTVPNPTGEDIILYEMETIRPFESHQIYRPLEPANMVGSDMSGPTIIVRGESVY 120
DB 61 kapntvpnptgediilylemetirpfishqlyrpdlepanmvgysdmspgtliivrgesvsv 120

QY 121 RFVNSGENTSPNSVHLHGSFSRAPFDGMAEDTQPGEXKYDYYPNROAARMLYHHDHAMS 180
DB 121 rfnsentspnsvhlhgsfsrapfdgmaedttqpgexkydyypnrgaamllyhdhams 180

QY 121 rfnsentspnsvhlhgsfsrapfdgmaedttqpgexkydyypnrgaamllyhdhams 180
DB 121 rfnsentspnsvhlhgsfsrapfdgmaedttqpgexkydyypnrgaamllyhdhams 180

QY 181 ITAENAYMGOAGVYMIQDPAEDALNPSGYGEFDPVLVITAKRYNADGTLFSTNGEVSSF 240
DB 181 itaenaymgoagvymiqdpaedalnpsgygefdpvlvltakrynadgtlfstngesvssf 240

QY 241 WGDVIOVNGOPMMLNVOPRKYRFRFLNAVSRFALYATSDSETRLPFGVIADGGL 300
DB 241 wgdviovngopmmlnvoprkylfrflnaavsrfsalylatsedsetrlpfqviadggll 300

QY 301 LGGPVDTDLIYISMAERWEVVIDFTFAGOSIDIRNLPGADGLGVPPEFNDTKVRFVY 360
DB 301 lggpvdtdcliysmaerwevidfstfagsidirlnpgadglgvpefndtkvmrfv 360

QY 361 DEVLESPTSEVPANLRDVPFPEGGMWDPANPTDDETFTEGRANGWTINGVTFSDVENR 420
DB 361 devlesptsevpantldvprfpeggmwdpanptdettftgrangwtlingvtfdsvenr 420

QY 421 LARNVROTVETIRLNNNGMTHPRVHILVDFRVLRSRARGVEPEYEAAGLKDVIWLAR 480
DB 421 larnvrotvvetirlnnnngmthprvhilvdfrrvlsrargvepeyeadgllkdviwlar 480

QY 481 RIEVYVEAHYAPFPYGLHCHNLIEHDDHMAAFNVTVLGIDGYVTEFIDMEPLMRP 540
DB 481 revvyveahyapfpyglhchnliehddhmaafnvtlvgidgyvtefidmepmlrp 540

QY 541 RPPLLGEFENGSGDSSELAITTDRIQEMASFNYPQAQDDDAEE 583
DB 541 rppllgefengsgdsselaittdriqemasfnypaqadddaaee 583

RESULT 2

AAB20097

XX AAB20097 standard; Protein; 583 AA.

AC AAB20097;

XX 23-APR-2001 (first entry)

XX Stachybotrys chartarum phenol oxidase B.
DE phenol oxidizing enzyme; phenol oxidase B; spob gene; bleach;
XX pulp; paper; textile; detergent.

OS Stachybotrys chartarum.

OS US6168936-B1.

PN 02-JUN-2001.

PF 22-SEP-1999; 9905-0401476.

PR 22-SEP-1999; 9905-0401476.

PA (GEMV) GENENCOR INT INC.

PI Wang H;

DR WPI; 2001-136715/14.

DR N-PSDB; AAF30028, AAF20029.

PT New phenol oxidizing enzyme, also useful in the detergent, paper and
PT pulp, textile or food industries, especially in modifying the colour
PT associated with dyes and coloured compounds, as well as in anti-dye
PT transfer applications

PS Claim 1; Fig 2; 23pp; English.

CC The present sequence is that of Stachybotrys chartarum MDCU 38898
CC phenol oxidase B, as deduced from isolated genomic DNA (see
CC AAF30028). The invention provides phenol oxidizing enzymes such
CC as phenol oxidase B, or enzymes having at least 68% identity to
CC it, nucleic acids encoding them, expression vectors, filamentous
CC fungus and yeast cells, and methods for the recombinant
CC production of the phenol oxidizing enzymes. The enzymes are useful
CC for bleaching pulp and paper, fabric stains, and in detergent and
CC textile applications. They show optimal activity at pH range 5-11,
CC 7-10.5 or 8-10, and at 20-60 or 20-40 degree C.

XX Sequence 583 AA;

Query Match 100.0%; Score 3114; DB 22; Length 583;

Best Local Similarity 100.0%; Pred. No. 8.9e-275;

Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISOAIGAVAGLAVIGSSVDARSVAGRSTDMPSGLTKRQTLSPPLALYEVLPIPL 60
DB 1 misqalgaavalglavlgssvdarsvagrstdmgsqtkrqtlspplalylevplippl 60

QY 61 KAPNTVPNPTGEDIILYEMETIRPFESHQIYRPLEPANMVGSDMSGPTIIVRGESVY 120
DB 61 kapntvpnptgediilylemetirpfishqlyrpdlepanmvgysdmspgtliivrgesvsv 120

QY 121 RFVNSGENTSPNSVHLHGSFSRAPFDGMAEDTQPGEXKYDYYPNROAARMLYHHDHAMS 180
DB 121 rfnsentspnsvhlhgsfsrapfdgmaedttqpgexkydyypnrgaamllyhdhams 180

QY 181 ITAENAYMGOAGVYMIQDPAEDALNPSGYGEFDPVLVITAKRYNADGTLFSTNGEVSSF 240
DB 181 itaenaymgoagvymiqdpaedalnpsgygefdpvlvltakrynadgtlfstngesvssf 240

QY 241 WGDVIOVNGOPMMLNVOPRKYRFRFLNAVSRFALYATSDSETRLPFGVIADGGL 300
DB 241 wgdviovngopmmlnvoprkylfrflnaavsrfsalylatsedsetrlpfqviadggll 300

QY 301 LGGPVDTDLIYISMAERWEVVIDFTFAGOSIDIRNLPGADGLGVPPEFNDTKVRFVY 360
DB 301 lggpvdtdcliysmaerwevidfstfagsidirlnpgadglgvpefndtkvmrfv 360

QY 361 DEVLESPTSEVPANLRDVPFPEGGMWDPANPTDDETFTEGRANGWTINGVTFSDVENR 420

```

Db 361 devlespdtevpantldvfppeggnwdpanplddefltgrangvclngvltfsvendr 420
OY 421 LLRNPBDTVEIMRLNNSNGMTHPVHILVDFRVLRSRARGVEPEAGLKDVMVLAR 480
Db 421 LLNRPDTEIWRLENNsgwtphvnhlvdfvrlsrstargvepeaaqldkvvmlar 480
OY 481 REVVVYEAHAPFPGVYMLCHNLIHEDHDMMAAFNVTVLGDYGYNTTEFTDPEPLMRP 540
Db 481 revvyeahapfpgvymhcnhlhdedhmmaafnvtvlgdgygnyteftidpmeplmrp 540
OY 541 RPFLLGEFENGSGDFSELATIDRIQEMASFNPPYAQADDDAAEE 583
Db 541 rpfllgefengsgdfselatidriqemasfnpyaqaddaaee 583

RESULT 3
AAB81506 standard; Protein: 578 AA.
ID AAB81506;
XX
AC AAB81506;
XX
DT 18-JUN-2001 (first entry)
XX
DE Billirubin oxidase.
XX
KM Billirubin oxidase; phenol oxidizing enzyme; phenol oxidase B;
KW detergent; paper production; pulp production; textile; food industry;
KM bleaching.
XX
OS unidentified.
XX
XX WO200121748-A1.
XX
PD 29-MAR-2001.
XX
PF 06-SEP-2000; 2000MO-EP08840.
XX
PR 22-SEP-1999; 99EP-0203120.
XX
PA (UNITL ) UNILEVER NV.
PA (UNITL ) UNILEVER PLC.
PA (HIND-) HINDUSTAN LEVER LTD.
XX
PI Convents D, Doornik M, De Vries CH, Wang H;
XX
DR MPI; 2001-273462/28.
XX
PT New detergent compositions comprising a phenol oxidizing enzyme useful
PT in detergent or cleaning compositions, fiber treatment, processing,
PT finishing or production, paper and pulp production, or in starch
PT processing applications
XX
PS Example 4; Fig 4; 46pp; English.
XX
CC The present sequence was used for comparison with the stachybotrys
CC charitum phenol oxidizing enzyme. The invention relates to detergent
CC compositions comprising one or more surfactants and a phenol oxidizing
CC enzyme having at least 68% identity to the Stachybotrys charitum phenol
CC oxidizing enzyme. Phenol oxidizing enzymes may be used in the detergent,
CC paper, pulp, textile and food industries. They are used for preventing
CC the transfer of dyes in solution from one textile to another during
CC detergent washing, or in modifying the colour associated with dyes and
CC coloured compounds having different chemical structures, such as in pulp
CC and paper bleaching, bleaching the colour of stains on fabric and in
CC detergent and textile applications.
XX
SQ Sequence 578 AA:

```

Query Match 99.2%; Score 3090; DB 22: Length 578;
 Best Local Similarity 100.0%; Pred. No. 1.3e-272;
 Matches 578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MISQIGAVALLGLAVIGSSVDARSVAGRSNDMPSGLTTRKROTSPPLALYEVLPIPL 60
Db 1 misqigavallglavlgssvdarsvagrscdmpsgltkrtqlspplalyevplippl 60
OY 61 KAPNTPNPTGEDILTYEMEIRPFSSHQIYPDLEPAMWGDGSPPTIIVRGTSVV 120
Db 61 kapntpnptgelltyemairpfshqiyppdlepanmwgdgmsppllivrgtssvv 120
OY 121 RFVNSGENTSPNSVHLHGSSSRAPFDGWAEDTTPQGEYKYYPPNRQAARMLWHDHMS 180
Db 121 rfvnsgentspnsvhlhgsstsrappfdgwaedtlqpgykyppnrqaarmlwyhdhms 180
OY 181 ITLKNAYMGAGVYMIODPADALNLPSCGEPTDPLVLAKRKNACGTLFSTNGEVSST 240
Db 181 itaenaymgagvyymiqdpadalanlpsgygetdplvlakrknadgltfstngvssst 240
OY 241 WGDYIQVNGQPWPMLANOPKRYRFRFLNAAVSRSFALYLTASEDSETRLPFOVIAADGGL 300
Db 241 wgdviqyngqpwpmldnvpkryrfrflnaavsrstfalyltasedsetrlpfvyaadggl 300
OY 301 LEGPVDITLYISMAERWEVYIDESTFAGOSIDIRNLPGADGLGVEPEDNTDKVMRFV 360
Db 301 legpvditllyismaerwevylidestfagoslidirnlpgadglgvepedntdkvmrfv 360
OY 361 DEVLESPDTEVPANLDRDVPFPEGGMNDPANPDDERTFEGFRANGQWTINGVTFSDVENR 420
Db 361 devlespdtevpantldrdrvfppeggmndpanpddertfegfrangvclngvltfsvendr 420
OY 421 LLRNPBDTVEIMRLNNSNGMTHPVHILVDFRVLRSRARGVEPEAGLKDVMVLAR 480
Db 421 LLNRPDTEIWRLENNsgwtphvnhlvdfvrlsrstargvepeaaqldkvvmlar 480
OY 481 REVVVYEAHAPFPGVYMLCHNLIHEDHDMMAAFNVTVLGDYGYNTTEFTDPEPLMRP 540
Db 481 revvyeahapfpgvymhcnhlhdedhmmaafnvtvlgdgygnyteftidpmeplmrp 540
OY 541 RPFLLGEFENGSGDFSELATIDRIQEMASFNPPYAQADDD 578
Db 541 rpfllgefengsgdfselatidriqemasfnpyaqadd 578

```

```

RESULT 4
AAR40843 standard; Protein: 572 AA.
ID AAR40843;
XX
AC AAR40843;
XX
DT 24-FEB-1994 (first entry)
XX
DE Billirubin oxidase.
XX
KM Billirubin oxidase; analytical; BO: expression vector; PCR;
KW polymerase chain reaction.
XX
OS Myrothecium verrucaria.
XX
FH key
FH Reptide 1..38
FH FT /label= sig_peptide
FH FT 39..534
FH FT /label= mat_protein
XX
PN JP05199882-A.
XX
PD 10-AUG-1993.
XX
PF 24-JAN-1992; 92JP-0034126.
XX
PR 24-JAN-1992; 92JP-0034126.
XX
PA (AMAN ) AMANO PHARM KK.
XX

```

DR WPI; 1993-284681/36.
 DR N-PSDB; AAQ47790.
 XX Bilirubin oxidase prep. useful as an analytical enzyme - by
 PT culturing bilirubin oxidase in transformant culture
 XX
 PS Claim 1; Page 29-32; 32pp; Japanese.
 CC The sequence encodes bilirubin oxidase. The protein produced has a
 CC 38 amino acid signal peptide which is removed to give the mature
 CC protein (AAR40843).
 XX
 SQ Sequence 572 AA;

Query Match 64.5%; Score 2010; DB 14; Length 572;
 Best Local Similarity 65.4%; Pred. No. 3,7e-174;
 Matches 383; Conservative 71; Mismatches 110; Indels 22; Gaps 10;

QY 1 MISQAIGAVAGLAVIGSSVDARSVAGRSTDMPSG-LTRKQOTLSPPLALYEVLPPIPP 59
 DB 1 mfhltgaaalsl-lfnsaevagpvp--etspatghlfrvqalspqymftvplpipp 57
 QY 60 LKAPN-TVPNPTGEDIITYEMEIRPFSHQIYFDLEPANWVGDSGPGFTIIVPGSTES 118
 DB 58 vkprrltvtpvnggelwlyveveikpfbhqvypdlgsadivgydgmspgplfgvprgvev 117
 QY 119 VYFVNSGENTSPKSVHLHGSFRRAPEDDQAEEDTQPGEKDYKYYRROAKRLWYHNDH 338
 DB 118 vyfllmae--apnsvllhgsfrraafdgwaedtlepgstkdyyprrgacrlwvhdha 115
 QY 179 MSITAEANAYMGAGVYMIODPADALNPSGYGEFDIPLVLTAKRYNADGTFSTNGEVS 238
 DB 176 mhtlaenayrgagqlymltdpdaedlnlpsgygefdlmlstkylangnltvtngeln 235
 QY 239 SFNGDIYQVNGGQWMLNKPQKRYRERFLNAVSRSAFLALASDSESETLPPQVIAADG 238
 DB 236 sfngdvlhvnvgqwpkknveprkyrriflidaavsrstfgyldatdrlpikvlasds 235
 QY 299 GLEGEVDVDTLXISMAERWEVVIDSTFAGOSIDIRNLPGA-DGLGVEPEFNTDKVMR 357
 DB 296 gllhepdaelslyismaeyevvfdtsdygktelelnlgsyggjctdcdydnctkvmr 355
 QY 358 FVVDVELESPTSEVPANLRDVPFPEGGMNDPMPD--ETFTFGRANGQMTINGVTFES 415
 DB 356 fvaadttcgdtcvsvpnanlrtdvpf-----sptnhrgrfgrtqptctwlngvafa 407
 QY 416 DVENRLLRNVPRTVETWRLNENSGMTWHPVHILVDFRLSRST--AGVEPYEAGL 472
 DB 408 dvgnrllanvpvgltveawelnaagngwclpilhldfkvlsrtsgnatltmpe-sgl 466
 QY 473 KDVVNLAREVVYVEAHYAPFPGVYMLHCHNLTHEDHDMMAAFNVVLADGYNTEFID 532
 DB 467 kdvvnlgrrctvveahyepfpgvymfchcnllhedhdmmaafnatvlpolygnatvfv 526
 QY 533 PNEPLRRPPELIGEFENGSGDFSELATDRIDEMASFNPNYQADD 578
 DB 527 pneeelwargpyelgefengsgdfsvavteridqtaeyrpyaade 572

RESULT 5
 AAF69204
 ID AAF69204 standard; Protein: 602 AA.
 AC
 XX
 XX AAF69204;
 DT 30-MAY-2000 (first entry)
 XX
 XX Amino acid sequence of a phenol oxidizing enzyme.
 KW Phenol oxidizing enzyme; fungus; redox reaction; detergent; bleaching;
 KW fabric; pulp; paper; decolourisation; plant-derived food product;
 KW coloured compound; porphyrin; tannin; polyphenol; carotenoid;

KW anthocyanin; Mallard reaction product.
 XX
 XX Acremonium murorum.
 OS
 XX WO200005349-A1.
 PN
 XX
 XX 03-FEB-2000.
 PD
 XX
 XX 13-JUL-1999; 99WO-EP04922.
 PF
 XX
 XX 21-JUL-1998; 98EP-0202454.
 PR
 XX
 XX (UNIL) UNILEVER NV.
 PA (UNIL) UNILEVER PLC.
 PA (HIND-) HINDUSTAN LEVER LTD.
 XX
 PI Convents D, Goka RJ, Van Der Heiden M, Swarthoff T, Verrips CT;
 DR
 XX WPI; 2000-195101/17.
 DR N-PSDB; AAZ61243.
 XX
 XX Phenol-oxidizing enzyme from Acremonium, used in detergent compositions
 PT for bleaching stains on fabrics -
 XX
 XX Claim 5; Page 41-43; 45pp; English.
 CC
 CC The present sequence represents a phenol oxidizing enzyme from the
 CC fungus Acremonium murorum. The enzyme has the CBS accession number
 CC 15772. The enzyme catalyses redox reactions and is specific for
 CC molecular oxygen as the electron acceptor. The phenol oxidizing enzyme
 CC is specifically used in detergents for bleaching stains on fabrics,
 CC but also for bleaching pulp and paper and for decolourisation of
 CC plant-derived food products. The enzyme has a pH optimum in the
 CC alkaline to neutral range and can bleach a wide variety of coloured
 CC compounds, e.g. porphyrins, tannins, polyphenols, carotenoids,
 CC anthocyanins and Mallard reaction products.
 CC
 SQ Sequence 602 AA;

Query Match 58.7%; Score 1829; DB 21; Length 602;
 Best Local Similarity 58.7%; Pred. No. 1.3e-157;
 Matches 352; Conservative 83; Mismatches 131; Indels 34; Gaps 7;

QY 5 AIGAVAGLAVIGSSV-----DARSVAGRST-----MPSGLTRKQOTLS 45
 DB 6 alrelalylslkgaqampkfeldlpeeeaaalaavedpandlqrrspndlqrrspls 65
 QY 46 PPLALYEVLPITPLKAP-NTVPNPTGEDIITYEMEIRPFSHQIYFDLEPANWVGDSG 104
 DB 66 paytlfgaplsiprvkeplftvnpnynggeidyelkhnfseqvfpdlqpadlvgydgl 125
 QY 105 SPGPITIVPRTGSESVRFVNSGENTSPNSVHLHGSFRRAPEDDQAEEDTQPGEKDYK 164
 DB 126 spgptlfigekgrsvrftvka--twesslhlngsrtapwgaevvtnpgeykdyyp 183
 QY 165 NROAARMLWYHNDHAMSITAEANAYMGAGVYMIODPADALNPSGYGEFDIPLVLTAKRY 224
 DB 184 naqagrfswyhdamhetaenayngaggyllndpedsiglpssygeydlpillskgy 243
 QY 225 NADGTLESTNGEVSFWGDIYQVNGQWMLNKPQKRYRERFLNAVSRSAFLALASDSE 284
 DB 244 nsdgtlftskgetqslwgdlygvnypwpyfdvpekrkyrrllidaavsrstfgyldad 303
 QY 285 SEFRLPFOVIAADGGLLEGVDVDTLXISMAERWEVVIDSTFAGOSIDIRNLPGADGIG 344
 DB 304 edrlrpfqvlaasdgllleevtsklvlsaeyellidsfegktelempavayg 363
 QY 345 VEPEPDNTKVMRF-VVDEYLESPTSEVPANLRDVPFPEGGMNDPMPDDETFTEFGRA 403
 DB 364 lervnyddctkvmrfvnaegplspdsctvsplrtvdpfss-----tsltldisffart 418
 QY 404 NGQWTINGVTFSDVENRLLRNVPRTVETWRLNENSGMTWHPVHILVDFRLSR----- 458

Query Match	55.18;	Score 1716.5;	DB 20;	Length 594;
Best Local Similarity	61.08;	Pred. No. 2.2e-147;		
Matches 330;	Conservative	67;	Mismatches 125;	Indels 19;
				Gaps 9;
OY	45	SPP-L-ALYEVLPRIPLKAPN-TVPNPNNGEILYEMERFSHOIYEDLEPANNGYD	102	
DB	57	sppynlllynaalrppkvkpkmlitcnrvtkgdwlyelakrfqarilypcllpatclvgyd	116	
OY	103	GMSPGFTIIVPRTGSVSVREVNSEGTSPNSVHLHGFSFRAAFDGAEDDTTQGEYKDY	162	
DB	117	gmsspftcfvnpqetclvrfina--tlvenshllhgsprapfdgvaedvtfpgeykdyy	174	
OY	163	YNNRQAAKRLWHLHDHAMSTTAENAYWGAGVYMIQDPAADALNLPBGCEFPRIPLYLTK	222	
DB	175	fnysasarllywhdhafmklaenayfkgagaylndeaadalglpsygefdlprllltak	234	
OY	223	RYNADGTLFSTGEVSSFEWGDYIQVNGQPMILNVPKRYRFRFLNAAVSRSFALYLATS	282	
DB	235	ynnadgltlstegeqddlwgdvlhvnqgwpflnqprytrfrflnaavstewllylvrt	294	
OY	283	EDSETRLPPOVLAADGGLLEGVDYDTLYLSAERREVVYIDSTFRGGSIDIRNLPGADG	342	
DB	295	asnpvnrplrfqvlasdagllqapvqtsnlylvaeryellidftntfagqldrnvaetnd	354	
OY	343	LGVEEFNDTDMVMEFVYDE-VLESBDTSEVPANLRDVPF--EGGMDDPANPTDDET	398	
DB	355	vgdedeayartlemvfsvsgtve--dsaqpsrltrdvpfrphkeg---pa---dkhf	404	
OY	399	TFGRANGQFTINGVFFSDVENLLRNVPRDYELIMRLNNSGWTPIVYIHLVDFRVLRS	458	
DB	405	kfeznghyhlndvfyadnervlakpeltgvevelenssgswshpvhhlhvdflklkr	464	
OY	453	STARG-VEVEFAAGKDDVYLARREVVYEAHYAPPGVYMLCHNLIEDHDMMAAFNV	517	
DB	465	tggrqvmypesaaglkdvwlgrgtcltleahyqpwtygmwhchllhcdndmafvrv	524	
OY	518	TVLGDYGVYVTEFIDMEPLMRPRPFLGCEFGSGDSESLAITDRIQEMASFNPYAQAD	577	
DB	525	tameekyglqedfedpmpkkravpynrnfharagnfaesitarvglaaeqeynrnd	584	
OY	578	D 578		
DB	585	e 585		
RESULT 11				
AA39992				
ID	AA39992	standard; Protein; 594 AA.		
AC	AA39992;			
DT	16-DEC-1999	(first entry)		
DE	Stachybotrys phenol oxidase protein sequence.			
KW	Phenol oxidase; enzyme; coloured compound; dye transfer prevention;			
OS	fabric washing; stain bleaching; anti-dye transfer; detergent.			
PM	Stachybotrys chartarum.			
PD	WO9949020-A2.			
PF	30-SEP-1999.			
PR	23-MAR-1999;	99WO-US06327.		
PR	24-MAR-1998;	98US-0046969.		
PR	22-DEC-1998;	98US-0218702.		
PR	22-MAR-1999;	99US-0273957.		

[illegible]

QY 496 VYMLCHNLHEDHMAAFNV 517
1: ||| | | | | |
Db 451 qvyhchylehsslgmaqlev 472

RESULT 15

AAAG0799
ID AAAG0799 standard; protein; 511 AA.

AC AAAG0799;

DT 26-SEP-2001 (first entry)

DE C glutamylum protein fragment SEQ ID NO: 4553.

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
organic acid synthesis.

OS Corynebacterium glutamicum.

PN EP1108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-0127688.

PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

PA (KYOW) KYOMA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR MPI; 2001-376931/40.

DR N-PSDB; AAH66018.

PT Novel polynucleotides derived from Coryneform bacteria, for identifying
mutation point of a gene, measuring expression of a gene, analysing
expression profile or pattern of a gene and identifying homologous gene

PS Claim 17; SEQ ID NO: 4553; 246pp + Sequence Listing; English.

CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum; these
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from Coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX Sequence 511 AA;

Query Match 16.2%; Score 504; DB 22; Length 511;

Best Local Similarity 29.5%; Pred. No. 5.9e-37;
Matches 157; Conservative 72; Mismatches 232; Indels 72; Gaps 18;

QY 7 GAVALGLAVIGG-----SSVDARSVAGRSTDPGSLTKROTLSPLALYEVPLPIPL 60

Db 12 gagvlatvgaqyivacssdvrrygggpr-----tlpippa 49

QY 61 KAPVVPNPTGEDIYEMERFSSHQIYPDLEPANMVGVDGMSGPTIIVPGTESVY 120

Db 50 dl-----gtregsav-hfaeaqtgesqllpdytktw-gfngtlhlppllvkkgddvhv 102
QY 121 RFVNSGENTSPNSVHLHGSPSRAPFDGMAEDTTPGEGYKRYYPNRGAAMLYHDIAMS 180
Db 103 dvinldemf--lvwhgmklpalaaggphspilpgqtwspwlvandaatlwyhphltg 160
QY 181 ITAENAYMGAGVYMIIDPAEDALNLPSCYGEFDIPVLTKAKRYNADGTLESTNGEVSSF 240
Db 161 ltglhayrtaagmliavedatdkldipreyvddipvlmdhftledgsidedlpolgl 220
QY 241 WGDVIOVNGQPPWMLNVQPKRYRFRFLNAAVSRSEFALYLATSEDESTRLPFOYIADGGL 300
Db 221 lgtprangitnahfdatttrvfrvinsgmrfynlaf-----sdtr-lfyjaasgl 274
QY 301 LEGVVDFTLYISMAEWEVVIDESTFAGSIDRNLRGADGLVE-----PERDNTDKV 355
Db 275 ldepqdrtlaigpgetweivele--pgeadvlesvgfednygvpddefvpdfigmsdf 332
QY 356 MRFVDEVLSPPTSEVPANLRDVPPECGNMPDANPTDDETFEGRANGQWTINGVTF 415
Db 333 qltlitg--psdaaagpa-----lpgylvkstep-dvidatertf-----lmtfsin 378
QY 416 DVENRLLR--NVPRTVEIWRLENNNSGWTHPVHHLVDFRVLRSRTANGVEPYEAGL 472
Db 379 dlqmdmqvrdvldhdqpevwlvclndsdwphmfhvndarfkvl-kfegtdvelfn-dgw 436
QY 473 KDYVWL--ARREVVYEAHAYRPPGYYMLCHNLHEDHMAAFNVYGLD 522
Db 437 kdtvglppgatatlavelfgnyppqpymyhcimlyhedgmmgqfvlvepgd 489

Search completed: October 3, 2002, 10:56:39
Job time: 143 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:54:16 ; Search time 34.9 seconds

(without alignments)
1605.160 Million cell updates/sec

Title: US-09-656-640A-2

Perfect score: 3114

Sequence: 1 MISQALGAVLGLAVIGSS.....IQEMASFNRYAQADDAAE 583

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2010	64.5	572	2 B48521	bilirubin oxidase
2	731.5	23.5	513	2 F69604	spore coat protein
3	489.5	15.7	568	2 B86364	hypothetical prote
4	460	14.8	533	2 AC0414	probable exported
5	442	14.2	527	2 C70397	periplasmic cell d
6	431	13.8	591	2 G96734	spore coat protein
7	416.5	13.4	494	2 AC3582	probable blue-copp
8	413.5	13.3	516	2 G90644	hypothetical prote
9	413.5	13.3	516	2 G85495	hypothetical prote
10	412.5	13.2	516	2 C64735	probable copper-bl
11	402.5	12.9	536	2 AF0523	probable copper-bl
12	339.5	10.9	474	2 AD0083	probable cell divi
13	322	10.3	470	2 E91116	suppressor of fsti
14	322	10.3	470	2 E85961	suppressor of fsti
15	319	10.2	470	2 G65088	suffi protein precu
16	310	10.0	470	2 AD0888	Suffi protein (impo
17	288.5	9.3	1662	2 T18540	mofa protein (impo
18	287.5	9.2	513	2 G81298	probable periplasm
19	259.5	8.3	463	2 G81375	probable metallo-o
20	256	8.2	500	2 B83910	hypothetical prote
21	247.5	7.9	622	2 S56214	probable membrane
22	247	7.9	311	2 H64157	suffi protein homol
23	246	7.9	504	2 F70813	hypothetical prote
24	218.5	7.0	721	2 H82528	L-ascorbate oxidas
25	199.5	6.4	551	2 T02752	probable laccase (
26	199	6.4	529	2 S68120	laccase (EC 1.10.3
27	196.5	6.3	622	2 S62580	probable multicopp
28	196.5	6.3	632	2 F83387	copper resistance
29	193.5	6.2	635	2 A36868	copa homolog - Xan

30	193	6.2	529	2 S49120	laccase (EC 1.10.3
31	192	6.2	619	1 KSNCTO	laccase (EC 1.10.3
32	191	6.1	580	2 F84828	probable laccase (
33	190.5	6.1	557	2 JC5229	laccase (EC 1.10.3
34	190	6.1	636	2 A55428	ferrioxalase precu
35	188.5	6.1	533	2 S62371	laccase (EC 1.10.3
36	188.5	6.1	570	2 H87368	copper-binding pro
37	188	6.0	548	2 S18746	laccase (EC 1.10.3
38	188	6.0	572	2 S68119	laccase (EC 1.10.3
39	187.5	6.0	527	2 JC5357	laccase (EC 1.10.3
40	187.5	6.0	621	2 S72493	laccase (EC 1.10.3
41	186.5	6.0	520	2 S59533	laccase (EC 1.10.3
42	183.5	5.9	520	2 JC5356	laccase (EC 1.10.3
43	183	5.9	619	1 KSNCT	laccase (EC 1.10.3
44	182.5	5.9	587	1 KSKVAO	L-ascorbate oxidas
45	181	5.8	554	2 T01240	laccase (EC 1.10.3

ALIGNMENTS

RESULT 1

B48521
bilirubin oxidase (EC 1.3.3.5) - fungus (Myrothecium verrucaria)

C/Species: Myrothecium verrucaria

C/Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000

C/Accession: B48521; A48521

R/Kikeda, S.; Ando, K.; Kaji, H.; Inoue, T.; Murao, S.; Takeuchi, K.; Samejima, T.

J. Biol. Chem. 268, 18801-18809, 1993

A/Title: Molecular cloning of the gene for bilirubin oxidase from Myrothecium verruca

A/Reference number: A48521; MUID:93366794

A/Accession: B48521

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-572 <KO11>

A/Cross-references: GB:D14081; NID:9436236; PIDN:BA03166.1; PID:9456710

A/Note: sequence extracted from NCBI backbone (NCBIN:136730, NCBIP:136732)

A/Accession: A48521

A/Status: preliminary

A/Molecule type: mRNA; protein

A/Residues: 1-572 <KO12>

A/Cross-references: GB:D14081; NID:9436236; PIDN:BA03166.1; PID:9456710

A/Note: sequence extracted from NCBI backbone (NCBIN:136728, NCBIP:136729)

C/Keywords: oxidoreductase

Query Match	Score	DB 2;	Length	572;
Best Local Similarity	64.58;	Pred. No. 1	1e-135;	
Matches	383;	Conservative	71;	Mismatches 110; Indels 22; Gaps 10;
QY	1	MISQALGAVLGLAVIGSSVDARSVAGRTDMSG-LTRKROLOSLPLALYEPPLIPP	59	
DB	1	MEKHTLGAALSL-LENSMAVQASVPV-ETSPATGHLFRVAISQYPPFTVPLIPP	57	
QY	60	LKAPN-TVPNPTNGEDILLYEMEIRPFSHOIYPLLEPANNVGYGMSPPFTIYPRCTES	118	
DB	58	VKQRLTVTPNVNQDELIMYEVEIKPFTHOYYPDLGADLVGIDGMSGPGFTYQVPRVET	117	
QY	119	VVRFVNSGENTSPNSVHLHGSFSRAPPDGMEDTTOGEXKDYYPNROARMLMYDHA	178	
DB	118	VVRFINNAE--APRSVHLHGSFSRAAPDGMEDITEGFSFDYYPNROARMLMYDHA	175	
QY	179	MSITAEAAVYMGQGVYTIODPAEDALMLPSGYGEFDIPLVLYAKRYADGLSTNGEVS	238	
DB	176	MHTIAENAYRGQAGVLTMLTPADBALMLPSGYGEFDIPLVLYAKRYADGLSTNGEVS	235	
QY	239	SFMGDVYQVNGQPPMNYOPRKYRFRFLNAAVSRFAVLATSEDETPRLPVOYVIAAG	298	
DB	236	SFMGDVYHVGQPPMNYOPRKYRFRFLNAAVSRSGLFADTIDTRLPVYVIAAGS	295	
QY	299	GLLEGVDTDTLYTSMARREYVYIDFSTFAGQSIDINLPQA-DGLGVEPEFDNTDKVMR	357	
DB	296	GLLEHFDTSILYTSMAERREYVYIDFSDYAGKTIELNLGSGIGIGTDTDYDNTDKVMR	355	


```

Query Match 14.8% Score 460 DB 2 Length 533
Best Local Similarity 28.3% Pred No. 56-25;
Matches 163; Conservative 62; Mismatches 210; Indels 140; Gaps 22;

QY 28 GASTMPGSLIKRQQLSPFLALVEPLPIPLKAPNTVPNPNGTGDILYEIMEIRFSH 87
Db 14 GATSLP-LMSRAA-----LADESPILPIPLLPDA-----KG 46
QY 88 QIYPLPEAPNW-----GYDGMSPGPTIIVPRGTSYVRFVNGSENTSPNSVHL 137
Db 47 KNLNLIGTGSYVWLSTATQWTGNGNLLGPAIRLORGAQVNTDITNALPEAT--TVHMH 104
QY 138 GFSFAPFPGMAEDTTPQPEYKDYVYPNBOAARMLYHHHNASITRNENAYMOAGVYMQ 197
Db 105 GLEIRGEVDCGQALIDPEAKRQYTFVAYEQPATCCFHHHTHSKTDHQVAKMDGLGLVID 164
QY 198 DPAEDALNIPSGYGEDPIVYLTAKRYNADGTLFNGEVS-----SFGGVIOV 247
Db 138 DPAEDALNIPSGYGEDPIVYLTAKRYNADGTLFNGEVS-----SFGGVIOV 247

```

Db 165 DSDSETLPKQKMGVDDIPVLLDQK-----LDDKHGQVYQYLDVMTAAAGWFGMDMLT 217

Qy 248 NGCPMPMLNVOPRKI-RRFLNLAASRSFALYATLSEDSFTRLPFOVIADGCLLEGVVD 306

Db 218 NGVPYPO-QITPRGVNRLRLINGCARSILNALSDCR-----PMYVLSAGGLAEFVY 270

Qy 307 TDTLYISMAEREEVVIDSTFAGOSIDIRNLPGAD-GIGVEPPEPDNDKVARF-----VV 360

Db 271 VRELITLNGEREFEVLD--TRDQOSLDLVTLPTVMQMTLAP-FDQPLPLRIQPSLAIG 327

Qy 361 DEVLESPDTSEVPANLRDVP-----FP 382

Db 328 SQVYL--PESLVIYIPELADYVTGYQERWFOPLMNDPKLMDLGMALVARYGMKAMAGMNNHG 385

Qy 383 EGGNNDPAMPD-----DEFTYGRANGQMTINGVTFSDVENRLRNVP 426

Db 386 DMGADHDENRDPMSGKKKKDHGTMGAPAFNFSSHAN--RINGKAFSMTPEAF--DAK 440

Qy 427 RDTVITWLENNNSNGWTPRVHILVDPFVLSRSTARGVEPYE-AAAGLKDVYWL--ARREV 483

Db 441 QCKYKWTITSGEDMLHPFHVHGTFPRIL--TENGKPPAEHRGMKWDIVVEGARSEI 497

Qy 484 VYVEAHVAPFPGVYVWLHCNLIHEHDHMMAPNVY 518

Db 498 LVRFNYLAPASTPYVAHCHLLEHEDTGMGLGFTYS 532

RESULT 5

C70397

periplasmic cell division protein (SufI) - Aquifex aeolicus

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999

C:Accession: C70397

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; M01D:98196666

A:Accession: C70397

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-527 <AOF>

A:Cross-references: GB:AE000724; NID:92983585; PIDN:AAC07157.1; PID:92983586; GB:AE000

A:Experimental source: strain VFS

C:Genetics:

C:Gene: sufi

Query Match 14.28; Score 442; DB 2; Length 527;

Best Local Similarity 28.7%; Pred. NO. 9,4e-24;

Matches 165; Conservative 77; Mismatches 204; Indels 128; Gaps 29;

Qy 10 ALGLAVIGGS--SYDARSVAGRSTDMPSGLTKRGTQLSPLALALEVLPPLKAPN--- 64

Db 14 ALGFSVGLSLILSCGGGGTSSSGGGTSLKSKSLNIPGYLF-----PDGQR 62

Qy 65 -----TVENPNTGCEIDILYEMEIFRFSHQIYFDLEPAMVGYDGMSPCPTIIV 113

Db 63 VSIATAKWTTLVIGKST--DMLYEID-----NEYNPVTLFR 98

Qy 114 RGTESVAVRFV--NSGENTSPNSVHLHGSFSAFP--DCMAEDTTPQGE--YKDYVYRNQ 167

Db 99 KGFQFSADFVNNSSGD--SLIHWHG--FRAPMKSDGHPYAAVADGETYSYDPTIIDRS 153

Qy 168 AARMMLVYDHAMSLTAENAKYMGAGVYMIQDPADALN--LPSTGYGEFDIPLYVTAKRYN 225

Db 154 GT--YFNPHPHGRGRGYQVYVYGLAGMIIEDEDENIKQALDLEYGIDIPLIQDTPFD 211

Qy 226 ADGTL--ESTNGEVSFWMGVIOVNGCPMPMLNVOPRKIRYRFLNLAASRSFALYATLSE 284

Db 212 SSGQGVYVNPQMNM-GFMQDITLVNLTPNPIYMDVSRKITYRRRIILNGSNAPRYRLALLRG 270

Qy 285 SETRLPFOVIADGCLLEGVDDTPLYISMAEREEVIDF-STFAGOSIDIRNLPG--- 339

```

Db      271  ---RRRFVVIGVEGSLDLTPKPEVNEILVAPGERIDILVDFRDSVNDYIKILYNPHNLIG 327
QY      340  -----ADGLGVEPEF-----DNTD-RVMRPVDEVLSEPTGSVPANLRV 379
Db      328  MGMIGMRGKMGKERMGCMGCMGNMMDGADNSEEFEVMEERYK--DSAYDKSIPLRLSV 385
QY      360  PPEEGGNDDPANFTDD---EFTTFRANGQWTINGVTF-----SDVEN-RLLRNVPRD 428
Db      366  -----TPIN-TGCAOVQRIITLGMRRVFTINGETWEDGYANPODINNPFVLEQUNNG 436
QY      429  TVEIWRLENNSGMTHPYHIHLVDPRVLSRSTARVEPEEA--GLKDVMLAREVVYV 486
Db      437  DVIILEEYVNNT-GMYHPMIHIGFOVQLERS---LGPLRATLDGMKDPVIVAPMETVRI 491
QY      487  E---AHYAAPFGVYGLHGNLTHEDHDMMAAAFNV 517
Db      492  AYDMSHPYNEHOIYLLCHLLEHDEGMAVNYRV 525

```

RESULT 6
G96734
spore coat protein-like protein, 24980-21957 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96734
R:Thelodopsis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Native 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mall, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719
A:Accession: G96734
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-591 <STO>
A:Cross-references: GB:AE005173; NID:56714313; PID:AAF26006.1; GSPDB:GN00141
C:Genetics:
A:Gene: F23N20.3
A:Map position: 1

	Query Match	13.8%;	Score 431;	DB 2;	Length 591;	
	Best Local Similarity	26.6%;	Pred. No. 6.8e-23;			
	Matches 143;	Conservative 75;	Mismatches 180;	Indels 140;	Gaps 22;	
QY	88 QIYPLLEPANNMGVYGMG-----PGPTTVPGRTGESVYRFPN-----	124				
Db	80 KFHRL -PATPEAAGTCSRATVCGPTTEAVYGDITVTKRNLPLHLITWDPDISPA	138				
QY	125 -SGENTSNSVHLGSGFSRAPDGMAE-----DTTQPGCYKDYVYENRQARRLMYH	175				
Db	139 IPKHGCIPTVYVHLHGCHIEPTSDGNADSMFTAGFKETGSKMTKTHYVKNKQPGMWVH	198				
QY	176 DHAMSTAEAMVGAQVYMIQ-DPAEDALN-PSGXEEDIDLVLTAKRYNADGTLF---	231				
Db	199 DHAAGLTIRNLLAGLLGSYILRHSSVESPLRPTG-REEDRDLVIFDRSFRDGSITYMA	257				
QY	232 -STNGEVSFSW-----GDVIOVNGQPMWPLNVQPRKYRFRFLNAVSRFALYLTSEDS	285				
Db	258 TGNNPETIRPQWQPEYFGDAIIYVNGKAMPRLTVRRKRYRPRITNASNAAPREFEESNG	314				
QY	286 ETRLRPFQVIADGGLLEGVPDITDLYISAEKMEVYVIDESTAGOSIDIRN---LPGADG	342				
Db	315 ---LDFIVVGSASAYLAKPVSTKSYLLAPSEIVDVLVDPSKSTSTALIANNAPYPYPSG	371				
QY	343 LGVEPFQNTDKVMRFVYVDEVLESPTLSVPLNKLADVPPPEGGMNDPAPT-----	393				

```

Db 372 -DPVTEENSKVAKKEIINNYKSE-VDTSIIPKLIEX-----PAHSTSTRTIYIA 419
Qy 394 -----DDETFPPGRANGOMTINGYTS-----DVENRLLRNPRTVTEIMKLENN 438
Db 420 MEEVYSSIDEPTHLY-----INGRPYANAYTEPKIQTSEKSLMTIMKEVEYINL 471
Qy 439 SNGMTHPVHILYDFRVLRSSTARCAVEPY-----EAALTKQVY----- 476
Db 472 TED-NHPLHILGLKEVLEQATLVKSEEFECMTKRNDAYKCEISYKARKNTAVYHER 530
Qy 477 -W-----TARREVVYEAANYA-----PRPGVYMLCHNLIHEDHDMMAAF 515
Db 531 GMRKVFEMKMPGHYTKILVRFYIHSNSYSRDAQDEGC-VYVHCHILDHEDNNMMRPF 587

```

```

RESULT      7
AC3582
Probable blue-copper protein yack precursor [imported] - Brucella melitensis (strain
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AC3582
R:DelVecchio, V.G., Kaparal, V., Redkar, R.J., Patra, G., Mijer, C., Los, T., Ivanov
: Mazur, M., Goldsman, E., Selkov, E., Elzei, F.H., Hagius, S., O'Callaghan, D., Lett
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitt
A:Reference number: AD3252; PMID:11756688
A:Accession: AC3582
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-494 <KUR>
A:Cross-references: GB:AE008918; PTD:ALU53822.1; PTD:g17984756; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110580
A:Map position: 11

```

```

Query Match 13.4%: Score 416.5: DB 2: Length 494:
Best Local Similarity 28.7%: Pred. No. 5.6e-22:
Matches 153: Conservative 61: Mismatches 205: Indels 117: Gaps 23:

QY 54 PLPIPLKAPNTVPNPNNGEDILLYEMEIRPEPSHOIYDPLLEPANNVGYDGMSPGPTIIVP 113
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 10 PLPLPLPEVPDA-----SG-----IYRLKAYQKRHS-FAKGSAAASAGINCALGLPLVRLM 59
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 114 RGTSEVVFPEVNSCENTSPNSVHLHGSESPAPDGAEDTTOGEXKDYVYVNRQAARMLM 173
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 60 SG-ESVTLSEVENAMD-EETTLTHMGLFPVSHLDGGPRHNVIAPAGAKWEPRKVAANQPSAFNM 117
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 174 YHDAHMSITAEANVYQAGVYVMIQDPAEDALMLPSGCGEFDIPLVLTAKRYNADGTLPEST 233
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 118 FHPHLHGHTARAHKGIAGLMTVIRGKAERGLPETYGVGDDIPLVLQDRIYLEGDAVAVP 177
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 234 N--GEVSFEMGVIOVNOQPMPLVWQPRKYFRFRLNAVSRSPALYLTSDSETRLPFE 291
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 178 DIMDLIHGFRGDMILVNGAIAPAEAVPAPAAVYLRLLNGAMNANFHIRPADGR-----PL 231
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 292 QVIAADGGLLEGPVYDITDLYTISMAERWEVVIDSTFAGQSIDI---RNLPGADGLGVEPE 348
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 232 LVIASDGGFISQPSVIEQTLTISPERGEYVLVDFSN--GEAVDLYVYGDNGSGDGL----- 284
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 349 FDNIDTKAKRFVYDVEYLES-----PQTSEVPA-----NLR----- 377
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 285 -----HLKRFYDVPALLEGRAVAKRPVSLDGPAPADDEKLSVQRRSFEDDERMAENMKMLMQ 339
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 378 ---DYVFPFGCGNMD-----PANPPTDETFPFGRANGWTINGYTF 414
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 340 PSSNPHASGDMDHDMKEMSGMACMDHMGSRSAADAGPALD-ALRSG--YOMALADKPE 395
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 415 S-----DVENRLLRLNVPKRDVETLWFRLENNNSNGTFRVHVLHILVFRVLTSTRAGVEPEYEA 470
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 396 DMERIDIVEAKL-----GSWEIWEI--TSREKAPHFHLHGASFRILSNAGKK--PPAHOT 445
   |||:||||: : :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

```

OY 471 GAKDYVWLAAREVYVE-----AHYAPPGYMYLCHNLHEDHDMMAFNVTY 519
 Db 446 GAKDTYLLIDGKAETLVHFDREARSHPP---MERCHLLEHEDVGMAQF-VTV 494

 RESULT 8
 G90644
 hypothetical protein ECs0127 [imported] - Escherichia coli (strain O157:H7, substrain R1
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 R:Accession: G90644
 R:Author: T. Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gassara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genoc
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: G90644
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-516 <HAY>
 A:Cross-References: GB:BA000007; PIDN:BA833550.1; PID:913359583; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain R1MD 0509952
 R:Genetics:
 A:Gene: ECs0127

 Query Match 13.3%; Score 413.5; DB 2; Length 516;
 Best Local Similarity 29.7%; Pred. No. 9.9e-22;
 Matches 141; Conservative 56; Mismatches 195; Indels 83; Gaps 19;

 OY 100 GIDGMSRPPIIYPRGIESYVRFNYSNGENTSPNSVHLHGFSRAFPDGAEDTTPGEEK 159
 Db 68 GYGNMLGPAVKLQGRKAVTVDIYN-QLPETTLHMHGLEVPGEVDGPGQIIPGGR 125

 OY 160 DYVYPRKQAAKMLVYHNSHAFENAMVGOAGVYMIODPAEDALNPSGEYEPDPLV 219
 Db 126 SVTLANDQPRATGCFHFNHOGKTRQVAMGLAGLVYIDDELTKMLPRQWGIIDDPVTV 185

 OY 220 TAKRYNADGTL---FSTNGEVSFPGDVTQYNGQWPMLNVPKRY-RREFNAAVSSEF 275
 Db 186 QPKKFSADGQIDYQLDVMTAAVGVFGDTLLTNGALYPO-HAARQMLRLRLNGCARSG 243

 OY 276 ALYLTASEDESEETLPROYVIAADGGLLEGSPVDTDTLYISMAERWEVYIDESTRAGSIDIR 335
 Db 244 -LNFAFASDNK---PLTVIASDGLLREPRVKVNEPLVMGGEFEVLVEVND--NKRFDLV 296

 OY 336 NLPQAD-GLGEVEPDNDTKVMRFVNDVELES---PDT-SEVPANLROVPEPGE----- 384
 Db 297 TLPVSQMGMAIAP-FDKPHFVMRIQPIAISASGALPDLTSLIPA---LPSTLEGLTVKRL 351

 OY 385 -----GNMDPANPTDDEFTEFGKANG 405
 Db 352 QLSMDMLDMQOMLMEKXYGDQAMVGMDSQMMCHMGCHNNHNN--HGKFDPLHAN- 408

 OY 406 QMTINGVTSDEVNRLIRNVPRTVEIWRLENNNGWTHPVHILVDFRVSSTARQVE 465
 Db 409 --KINQOAEF--MNKPFMAAKQGERWVIVSGVGMMLHPHIIHQTFRIISEN--GKP 461

 OY 466 P-YEAGLKDQVNLARR--EVVYVEAHYAPPGYMYLCHNLHEDHDMMAFNV 517
 Db 462 PAHRAGMKDTVKEGVNVEVLKFNHDAPEKRAVMAHCHLLEHEDTGKMLGETV 516

 RESULT 9
 G85495
 hypothetical protein yack [imported] - Escherichia coli (strain O157:H7, substrain EDL93
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: G85495
 C:Author: N.T. Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potluri, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7

A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85495
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-516 <STO>
A:Cross-references: GB:A0005174; NID:g12512835; PIDN:AAG54427.1; GSDB:GN00145; UWGP:
C:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yack

Query Match 13.3%; Score 413.5; DB:2; Length 516;
Best Local Similarity 29.7%; Pred. No. 9.9e-22;
Matches 141; Conservative 56; Mismatches 195; Indels 83; Gaps 19;

OY 100 GDGSGSPRTIIIPVPGTSVVREVNCGENTSPSNVHLHSSRAFPDQMAEDTPOPEYK 159
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 68 GYNGLMLBRPAVKQLQGSKAVTDIVN--QLTEETTLHMHGLEVREGVDGGPQGIIRPGKKR 125

OY 160 DYUYPNRQAARLMWYNDHAMSTTAENAYMGQAGUYMIODPAEDALNLPSGYGEEDIPLVL 219
 ||| - | - | - | - | - | - | - | - | - | - | - | - | - | -
Db 126 SVTLNVDPRAATCWPHRNQHNGKTGRQVALMGLAGLVIEDDELTKMLPKOMGIDDVPYIV 185

OY 220 TAKRRNAOGTL---ESTNGEVSSEFWGDVIQNMGOWPRMLNQPKKY-KFRELAAVSSSF 275
 |||::||| : | : | : | : | : | : | : | : | : | : | : | : | :
Db 186 QDKKESABGOIUDYOLDVTAAVGWDITLTNGALTPO-HAAPRCMLRLRLNGCNARS- 243

OY 276 ALYTISEDESETRLPFEOYLADGDLLEGVRPDPTLYISMAERWEVIDFSFGAGSIDIR 335
 ||| :: : | : | : | : | : | : | : | : | : | : | : | : | :
Db 244 -LNFTSDNR----PLYIASDGILLPERPVKNELPYLMGERFEVLVYND-NKPFDLV 296

OY 336 NLPGAD-GLGVEPERDNDDKWRFFVDEVLES---PDT-SEVPANLRDVPPEREG--- 384
 ||| : : : | : | : | : | : | : | : | : | : | : | : | :
Db 297 TLPIVQMGMATAP-FDKRPVMRIQRIASAGALPDDLSSLPA---LPSLGELYAKL 351

OY 385 -----GNMDPANPPDDETTFTRANG 405
 ||| : | : | : | : | : | : | : | : | : | : | : | : | :
Db 352 QLSDMPMLDMGMOMIMKERYGDAAYVGMDSQMGCHMGNNHN--HGKRFETHAN- 408

OY 406 GWITGVTFSDVENELLRNVRDVAIEIRLENSNGTHVPHINILVDFRSRPARGE 465
 ||| : | : | : | : | : | : | : | : | : | : | : | : | :
Db 409 --KINGQAFD-MNKPRFAAKGQTERENVVISGVGDMMLHPRHIGTORFILSEN---GKP 461

OY 466 P-YEAAGLDVVMIAARR--EVYVYEANYAPPRGVYMILCHNILIHEDNMMAFNV 517
 ||| : | : | : | : | : | : | : | : | : | : | : | : | :
Db 462 PAAHBAGMKDYFKVGENVSEVLVKFNHDAPKERAAMACHILLEHEDTGMMIGFTV 516

RESULT 10
C64735
Probable copper-binding protein yack - Escherichia coli
C.Species: Escherichia coli
C.Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text-change 08-Oct-1999
C.Accession: C64735; S45200
R.Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A.Reference number: A64720; MUID:97426617
A:Accession: C64735
A.Status: nucleic acid sequence not shown; translation not shown
A:Residues: 1-516 <BLAT>
A:Molecule type: DNA
A:Cross-references: GB:A0000121; GB:U00096; NID:g1786306; PIDN:AAC73234.1; PID:g1786303
A:Experimental source: strain K-12, substrain MG1655
R.Fujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45181
A:Accession: S45200
A:Molecule type: DNA
A:Residues: 1-463, 'LIARAQ', 470, 'IRLR' <EMBL>
A:Cross-references: EMBL:D26562; NID:g473770; PIDN:BBA0579.1; PID:d1006123; PID:g4737
A:Experimental source: Strain K-12, substrain WJ110


```

QY 250 QPWLPLNQPRKRYRFRFLNAAVSRSEFALYLTSEDSERLPEFOYLAAGGLLEGVPYDIT 309
Db 218 AQSPVEVSRGWRLRLNLNASNARYTLQLSDBG-----PLYVASDQGLPAPVAOQ 271
QY 310 LYISMAERWEVVIDF-----STFAGQSIDI-----RNLPGADGL-----GYEP 347
Db 272 LSLAAGERREVVIIIDMSQNEVSIITFAGESAGIMDRLRGLFEPSSILITVLTLKPTGLP 331
QY 348 EEDNDKX-MRFVYDEVLESPTSEVPANLRDVPPEEGGNMDPAFPTDEFTTERANQ 406
Db 332 LV--TDNLPMRLADQIIIEG-----SVIRSREQLDNLNPGIN-----GAI 370
QY 407 WTINGVFESDVENRLLRNVPDRDVEIMRLNENNSNGTAPVHIILHVDYFLSRSTARVEP 466
Db 371 WDMNV-----DYQ-----AQQGTWERIITHAD--MPQAFHILQGVSLYKSVNGAANA- 417
QY 467 YEAGLRKYVWL 478
Db 418 -EDRGWKDTAV 428

```

RESULT 13
E91116
suppressor of ftsi [imported] - Escherichia coli (strain O157:H7, substrain R1MD 050995)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence,revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: E91116
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhnara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
J. Bacteriol. 183, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic
A:Reference number: A99629; M01D:21156231; PMID:11258796
A:Accession: E91116
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-470 <HAY>
A:Cross-references: GB:BA000007; P1DN:BA837324.1; P1D:g13363373; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
C:Gene: Ecs901

[illegible]

```

Db 337 GLEPLV--TDSLPMLLPTEIMAGS-----PIRSRDISLGG--DPG-----363
Qy 403 ANGO-WTINGTFSFSDVEHRLLRNVPDVIETWIRLENNSGNTHPVHILHYDFRYLSTRA 461
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 334 INGQMDVNR1--DV-----TQOQGWEWRYTRADE---PQAFHLEGWYFQI---RNV 408
Qy 462 RGVERY-EAAGLKVAYWL---AREEVYVEVHYTFEP 494
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 409 NGAMFFPEDRGMDKIVWDVGQVELLYVFGQPSMAHP 445

```

RESULT 14
E85961

C:Species: Escherichia coli
C:Date: 16-Feb-2001
C:Accession: E85961
C:Accession: E85961
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apod
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206351
Z:Accession: E85961
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-470 <STO>
A:Cross-references: GB:AE005174; NID:q12517586; PIDN:MA658153.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: sufi

Query Match	10.3%	Score 322;	DB 2;	Length 470;
Best Local Similarity	25.3%;	Pred. No. 2.9e-15;		
Matches 131;	Conservative 68;	Mismatches 204;	Indels 114;	Gaps 21;

```

0Y 10 ALGLVAJGSSVADVASVGNKSTDMPSGLTKRQJOLSPPLALVEYVLPJLPLKADNPVNP 69
Db 11 ASGIMLCAGAPLKLASAGQ-----OOPLVPPL-----ES 42
0Y 70 NTGEDIILYEWEIREPFSHOIIPDLFANMVGVDGMSPEPTIIVRGTESVVRFNNGENT 129
Db 43 RRGQPLF--WTVOC-AHWSEFPCTRASVWGINGLYLPTLRWKGDD--VKLISNRJLT 96
0Y 130 SPNSVHLHGFSRAFPEDMAEDTTQPGCEKXYYPNROAEMLYHJHMASITAEAMYG 189
Db 97 ENVSMTVAGLOVPGPLMGCPARBMSPNADMARVLPJRNATLWYHANTPRTAOOYNG 156
0Y 190 QAGVYMIODPADEDALNLPSCYGEFDPILVLTAKRYMADGTLFSTINGEYSSWGVDIOYNG 249
Db 157 LAGMWLVEDEVSKEPLPIHNGYVDPEVLIODKRLDNNGTPEYNEPSCGSGVCGJTLTNG 216
0Y 250 QPMPLANTQPKRRYRRFRFJANVSRSFALYATSEDSERLPEOYATAGGLLEDPVDT 309
Db 217 VQSPVEVSRQWVRRLRLMASNSRKYQLQMSDGR-----PLHYSIDGGLPARVYSKQ 270
0Y 310 LYISKAERWEVVIDEST-----FAGQSIDIRNL-----PGADGL 343
Db 271 LSLAPGEERREILVMSNDGEVSIITCGEAMASIVDRIGCFEFPSSILVSTLYTLRP---T 326
0Y 344 GVEPEFDWTDVU-KRFVYDEVLESPTSEVPANLRDVPFFPGGMMDPANPIDDTEFFGR 402
Db 327 GLPLPV-TDSLPMRLPTEITMAGS-----PIRSBDISLGD-----DPe----- 363
0Y 403 ANGO-WTINGTFSDEVENRLRNVRPDTVEIMRLSNSMGWTHPVIHLYDFRVLSTRSTA 461
Db 364 INGOLMDVNRJ---DV-----TAQGTWERWYVRADE--PQAFHIEGVAFQJ---RNV 408
0Y 462 RGVERY-EAAGLKVWVL---ARREYVVEAHYAPFP 494
Db 409 NGAMFFEDRGWKDTVWVDQVEVLLVYFGQPSMAHP 445

```

RESULT 15
 sufi protein precursor - Escherichia coli
 G65088
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 08-Oct-1999
 C:Accession: G65088; S20461
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: G65088
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-470 <BLAT>
 A:Cross-references: GB:AE000384; GB:U00096; NID:q2367186; PID:AMC76053.1; PID:q1789394;
 A:Experimental source: strain K-12, substrain MG1655
 R:Coleman, J.
 Mol. Gen. Genet. 232, 295-303, 1992
 A:Title: Characterization of the Escherichia coli gene for 1-acyl-sn-glycerol-3-phosphat
 A:Reference number: S20459; MUID:92212294
 A:Accession: S20461
 A:Molecule type: DNA
 A:Residues: 1-89 <COL>
 A:Cross-references: GB:M63491; NID:q147296; PID:AAA24398.1; PID:q147299
 C:Genetics:
 A:Gene: sufi

Query Match 10 28; Score 319; DB 2; Length 470;
 Best Local Similarity 25.38; Pred. No. 4.8e-15;
 Matches 131; Conservative 67; Mismatches 205; Indels 114; Gaps 21;
 10 ALGLAVIGSSVDARSAGRTDMPGLTRQTLSPPLALYEVPLPPLKAPNTVPNP 69
 11 ASGINALCAGAVPLKASAAQ-----QQPLPVPPL-----ES 42
 70 NMGEDILYEMEIIPPSHOIYDLEPANVGYDGMSPGPTIIVRGTESVRFVNSGENT 129
 43 RRGQPLF--MTVQR-AHMSFTPGTRASVWGINGRYLGPTRVWKGD--VKLIYSNRLT 96
 130 SPNSVHLHGSFRAPFDGAEDTQCEYKDYYPNRQAARMLWYHDHAMSTIAENAYMG 189
 97 ENVSMIVAGLOVPGPLMGSPARMSPNADWAPVLPTRQNAATLWYHANTPNRTAQQVYNG 156
 190 QAGVYMIODPADALNLPISGYEFDIPLVLTAKRYNADGLTFSTNGEVSFWGDVIOVNG 249
 157 LAGMMLVEDEVSKSLPIPMHYGVDDPVIYIODKRLDNFGTPEYNERGSGGFVCDTLVNG 216
 250 QPWPMLNVOFRKYPFRFLNAAVSRSFALYLATSESETRLPFOVIYADGGLLEGVDTDT 309
 217 VQSPYVEVSRGWVRLRLNLSNSRRYLOMNDGR-----PLHVISGDGFLPAPVSVKQ 270
 310 LYISMAERWEVVIDFST-----FAGOSIDIRNL-----PGADGL 343
 271 LSLAPGERREIILVDSMGDEVISITGEEAASIVDRIRGFEPPSSILVSTLVLTLP-----T 326
 344 GVEPEFDNTDKY-MRFVYDEVLESPTSEVPANLRDVPPEGGNMDPANPTDDEFTFGR 402
 327 GLLPLV--TDSLPMRLLPTEIMAGS-----PIRSRDISLGD-----DPC-----*363
 403 ANGO-WTINGVTFSDEVENFLRNPRDIVEIMBLENNSGWTHPVHILVDFRVLRSSTA 461
 364 INGQLMDVARI--DV-----TAQGTWERWTVRADE--POAFHIEGVMEFI--RNV 408
 462 RGEVRY-EAAGLKDYYVWL--ARREVVYEAHYAPRP 494
 409 NGAMPPEPDGRGKMTVWVDQVELLYVEGQPSMAHP 445

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using SW model

Run on: October 3, 2002, 10:55:33 ; Search time 34.9 Seconds

(without alignments)
1574.874 Million cell updates/sec

Title: US-09-656-640A-4

Perfect score: 3067

Sequence: 1 MFKHTLGAALSLFLNSNAV.....AVTERIQMAEYPPYAAAE 572

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: / 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3067	100.0	572	B48521	bilirubin oxidase
2	713	23.2	513	F69604	spore coat protein
3	506	16.5	568	B86364	hypothetical prote
4	470	13.3	591	G96734	spore coat protein
5	437.5	13.3	533	AC0414	probable exported
6	403	13.1	494	AC3582	probable blue-copp
7	401	13.1	516	C64735	probable copper-bi
8	400.5	13.1	527	C70397	periplasmic cell d
9	398	13.0	516	G90644	hypothetical prote
10	398	13.0	516	G85495	hypothetical prote
11	398	13.0	536	AF0523	probable multicop
12	333	10.9	474	AD0083	probable cell divi
13	318.5	10.4	470	E91116	suppressor of ftsI
14	318.5	10.4	470	E91116	suppressor of ftsI
15	315.5	10.3	470	G65088	SuII protein precu
16	311	10.1	470	AD0888	SuII protein (impo
17	276.5	9.0	1662	T18540	mofta protein precu
18	276	9.0	513	G81298	probable periplasm
19	261	8.5	463	G83175	probable metallo-o
20	260	8.5	500	B83910	hypothetical prote
21	241.5	7.9	504	F70813	hypothetical prote
22	228	7.4	311	H64157	SuII protein homol
23	215	6.9	520	S59333	laccase (EC 1.10.3
24	212	6.9	520	JC5356	laccase (EC 1.10.3
25	209	6.8	557	JC5229	laccase (EC 1.10.3
26	207	6.7	527	JC5357	laccase (EC 1.10.3
27	205.5	6.7	520	A35883	laccase (EC 1.10.3
28	205.5	6.7	520	B35883	lignolytic pheno
29	205	6.7	609	KSPSCY	copper resistance

30	201.5	6.6	580	F84828	probable laccase (
31	199.5	6.5	548	S18746	laccase (EC 1.10.3
32	198	6.5	551	T02752	probable laccase (
33	197.5	6.4	529	S68120	laccase (EC 1.10.3
34	197.5	6.4	554	T01240	laccase (EC 1.10.3
35	197	6.4	624	A36962	laccase (EC 1.10.3
36	195	6.4	721	H82528	L-ascorbate oxidas
37	194	6.3	619	1 KSNCL0	laccase (EC 1.10.3
38	193.5	6.3	622	S56214	probable membrane
39	191	6.2	553	T45959	laccase-like prote
40	191	6.2	605	S52253	copper resistance
41	189.5	6.2	536	A55428	ferroxidase precu
42	189	6.2	533	S62371	laccase (EC 1.10.3
43	187.5	6.1	621	S72493	laccase (EC 1.10.3
44	185.5	6.0	529	S49120	laccase (EC 1.10.3
45	184.5	6.0	611	C82845	copper resistance

ALIGNMENTS

Query Match	100.0%	Score 3067	DB 2	Length 572
Best Local Similarity	100.0%	Pred. No. 1.8e-217		
Matches 572	Conservative	0	Mismatches 0	Indels 0
QY	1 MFKHTLGAALSLFLNSNAVQASPVETSPATGHLKRVAAQISPOVPMFTVPLPIPPVQ 60			
DB	1 MFKHTLGAALSLFLNSNAVQASPVETSPATGHLKRVAAQISPOVPMFTVPLPIPPVQ 60			
QY	61 PRLTVNPNVNGOEIWEYEIEIKFTHQVYPDLGSADLVGIDGMSPPPTFOVPGVETVVR 120			
DB	61 PRLTVNPNVNGOEIWEYEIEIKFTHQVYPDLGSADLVGIDGMSPPPTFOVPGVETVVR 120			
QY	121 FINNAEPNSVHLHGSFRAAFDGMADITEPGSFQDYVYPNROSARTLWYHDHAMHITA 180			
DB	121 FINNAEPNSVHLHGSFRAAFDGMADITEPGSFQDYVYPNROSARTLWYHDHAMHITA 180			
QY	181 ENAYRQAGLYMLTDPADALNLPSCYGEFDIPMLITSQYANGMLVTTNGELNSFWGD 240			
DB	181 ENAYRQAGLYMLTDPADALNLPSCYGEFDIPMLITSQYANGMLVTTNGELNSFWGD 240			
QY	241 VIVVNGQPMPEKFNVEPRKRYFRFLDAVSRSGFLYADPDADIDTRLPFKVIASDSGLLEH 300			
DB	241 VIVVNGQPMPEKFNVEPRKRYFRFLDAVSRSGFLYADPDADIDTRLPFKVIASDSGLLEH 300			
QY	301 PADTSLIYSMAERYVVFEDSDYAGKTIELRNLAGSGISGIGTDTDTYDNTDKVMRFVAD 360			
DB	301 PADTSLIYSMAERYVVFEDSDYAGKTIELRNLAGSGISGIGTDTDTYDNTDKVMRFVAD 360			

Query Match 23.2% Score 713; DB 2; Length 513;
Best Local Similarity 33.1% Pred. No.1,5e-44;
Matches 177; Conservative 77; Mismatches 179; Indels 102; Gaps 16;

49 FTVPLEIP---PVKQPRLTVTNPVNGGEIWEIKRFTTHQVYPPDGSGADLVGTDGMS 104

```

Db      6 FVDADRIDDLTKLPVQO-----SKEKUYEUVMEECTHOLHRLDPTRYLGMNGLE 55
QY      105 PGPFPQVGRGVEYVAFPTNN-----AEPASVHLHGSFSNAF 142
Db      56 FGPTEVYKRNENYVYKMMNNLPTSHFLPIDHTIHNSOCHREPEVYKTVYVHLHGYPBDS 115
QY      143 DGMAB-----DITEGSE--KDYVYRPNROSARTLWYHDHAMHITLAEVYAGAGLYMT 194
Db      116 DGYPEAWSEKPEOETGPRFKREYVYHYPNOOGAILMYHDHAMALTRYVYAGLVAYIHH 175
QY      135 DPAEALULPGSGYGEODRIYLTGKOTANENLWTTGCEL-----SFMGDVI 242
Db      176 DPKERKLWPS--DEYDPLILTDRTINEDOSLTYPPAPENPSPSLPNSIVAFQCEI 233
QY      243 HVNGQPMFKNVPERKRYRFLDAVBSFGLYADTDAIDTLRPFKYVIASDGLLEHFA 302
Db      234 LVNGKVVYVLEUEPRKYRFRVIMASNTRYNL-----SLDNGDFFQIOSDGLLPRSV 287
QY      303 DTSLLYJMAEBYEVVFPESDYACKTTELRLGSGIGIDPTGYDNTDKVMRVPVADDT 362
Db      268 KUNSSILPAERETDIIIDFTLVEGESILANSACCGDVPETDAN-----IMQFRVYKPL 343
QY      363 TQPDTSVVPANLRDVPSPPTTNPQFR-----FGRTGPTWTNGVAFADYQNR 412
Db      344 AQKDSRKRPKYL--ASYSVOYHERIQNIRTLKLAGTODEYGR--PVLLNNKKRWHD-- 366
QY      413 LLANVP--GVTEYRMELINAGNGWTHPIHILYDVKVYSRTSGNNAR-----457
Db      397 -VTEPRKGTETHEWISINPTRG-THPIHILVSRVLDIRDRFDIARVQOESGELSYTGPV 454
QY      456 TVMRYESGLKDYVWLGRRETYVVEAHVYAPFGYVIMFCHNLIHEDHDMMAMAFNAT 512
Db      455 PPPPEKWMKDTIQAHAGEVYLRATERYSGRYVWICHILHEDDMMAMRMDIT 509

```

```

RESULT      3
8B6364
hypothetical protein F19G10.5 - Arabidopsis thaliana
C|Species: Arabidopsis thaliana (mouse-ear cress)
C|Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C|Accession: B86364
R|Theologis, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A|Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Matli, R.; Marzita
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A|Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A|Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A|Reference number: AB6141; MUID:21016719
A|Accession: B86364
A>Status: preliminary
A|Molecule type: DNA
A|Residues: 1-568 <STO>
A|Cross-references: GB:AE005172; NID:92462832; PIDN:AAE72167.1; GSPDB:GN00141
C|Genetics:
A|Map position: 1

Query Match          16.5%; Score 506; DB 2; Length 568;
Best Local Similarity 29.3%; Pred. No. 2.8e-29;
Matches 168; Conservative 70; Mismatches 180; Indels 156; Gaps 27;

QY   48 METVLPPIPPVQOPRLITVNPNVNGGEIWIYEVEIKRPFTHQV-----YPDLGSADLV 98
    ||  ||              |||| : | : ||| : : : : : ||
Db   29 MEVDLDP-----DMPRLYGNSVHG-----IIKPSAQIQMFSTKKFHRDLPATPVF 76
    ||  ||              |||| : | : ||| : : : : : ||

QY   99 GYDGMS-----GGPTFOVPRGCVETVVRFIN-----NAEAPNSVH 132
    ||  ||              |||| : | : ||| : : : : : ||
Db   77 AY-GTSRSKAIVPGPIIEIVGVGDVIYVTNRNLPKSHILPMQPTISPAATPRKHGGLPTVVH 135

```

```

QY 133 LHGSESRRAFDGMADITEPGSKD-----YYPNRSARLTMYHDHAMHITAENA 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 LHGHIHEPTSDNA-DAMTAGFREREPKMTKTLHYEKKQDGNMWMYDHAMGLRLV 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 184 YGQAGLYLTDPA-EDALNPSGGEEDIPMLTNSKQYTAGN---NLVTNGELNSFW 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 LAGLVAGYLLRRHAAVSPQLPTG-DEFDRPLIEFRSFRKGSILYMNATGNPISHPQW 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 -----GDVTHVNGQWPFKRVNPERKREFRLDAVSRSGLYFADDAIDTLRPFVITS 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 254 QPEYFGDVIIVGKAMPRLNVRRRYRFRILNASNAFEKFEFSN-----GLDFTVGS 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 294 DSGLEHPADTSLITYISMAERYEVDFSDYAGKTIELNLGSLIGIGTIDY-----347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 308 DSAVLSKPMVTSILSPSEIYDVVDYFKSPSRIVLAN-----DAPYRPSGD 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 348 ---DNTDKVMRVVADDTQPTSVVPAHLRDPVPPSPPTTTPROFGRGTGPTWINGV 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 PVNEGKVMKELI-NNESDDTCTIPKTL--INPNADVSNALVTRY-----ISMVEYV 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 AFADYQNRLLAN-----VYV-----GTERWELINAGNGWHPHILHVDKVISRTS- 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 SNSDEPTHLVNGLYEADVETPKSGTTEVWEVNLTRD-NHPLHILGLRKYVEQTL 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 453 -----GNNM-----RTVMPYESGLKDVWL--GRETVEYEAH 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 LLAGLEEFKECKTKNDANKCOISKYARCKTAIVTAHEGKMNVPFMGCHYTRILIVRS 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 484 VA-----PF-----PGVYMFCHNLHEDHDM 506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 YIHTNASYFDPDTEPG-VYVHCHILHEDHDM 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 4
spore coat protein-like protein, 24980-21957 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G967734
R:Thelodios, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
  Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
  ansen, N.F.; Hughes, B.; Hultzar, L.
  Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
  C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Matzali,
  Rizzo, M.; Rooney, T.; Kowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
  ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719
A:Accession: G967734
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-591 <STC>
A:Cross-references: GB:AF005173; NID:g6714313; PIDN:AAF26006.1; GSPDB:GN00141
C:Genetics:
A:Gene: F23N20.3
A:Map position: 1

```

```

Query Match 15.3%; Score 470; DB 2; Length 591;
Best Local Similarity 29.6%; Pred. No. 1,3e-26;
Matches 156; Conservative 66; Mismatches 183; Indels 122; Gaps 24;
QY 91 DUGSADLVGYDGM-----PGPTQVPRGVETVAFIN-----N 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84 DLPATPVFAV-CTSKSATATVPPTIAVGVDTYVWRNHLPLHMLPMDPTSPAIPKH 142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 AAPAPSVNHHGFSRAAFGMAED-----ITEPGS---FKDYIYPRKQARILMYDHAM 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 143 GGIPTVVLHGHGHEPTSDGNSDWFETAGFKETGSKMTKTHYVVKQDPRGNMWDHAA 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 177 HTAENAYRGQAGLYMLT-DPAEDALNPSGGEEDIPMLTNSKQYTAGN---NLVTNN 231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 GLTRVALLGLGSLYLRHSSVESPLRPTG-REFRPLVIDFRSRKQGSIMNATGN 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 232 GELNSFW-----GDVTHVNGQWPFKRVNPERKREFRLDAVSRSGLYFADDAIDTRL 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 PIHPQOPEYFEDALIVGKAMPRLTVRRKRYRFRITNASNAFRFEFSN-----GL 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 287 PFVIVSDSGLEHPADTSLITYISMAERYEVDFSDYAGKTIELNLGSLIGIGTIDND 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 DFTVGSDSAYLAKPVSTKSVLAPSEIYDVLDVDFSKSTKATILNANAPYPSGDPTV 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 347 YDNTDKVMRVVADDTQPTSVVPAHLRDPVPPSPPTTTPROFGR-----GRTGPT 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 EENS-KVMFII-NYKSEVDTSIIPKTL--IEPRHNVSTSTRTIRIAMEVYSSIDEPT 431
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 399 WT-INGVAFADYQNRLLANP-VGIYE-----RRELINAGNGWHPHILHVDK 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 HLYINGLPY---NAVTEFPKIGTSEFSLSMLTIKQVWEVNLTRD-NHPLHILGLRK 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 447 V-----ISRTSGNNARTVMPYESGLKDVWL--GRETIV 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 VLEQTRALVKSSEPTIECKTRNDVAKCEISKYARCKTAIVTAHERGKKNVFKMPGHVTKI 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 479 VVEAHY-----APPGVYMFCHNLHEDHMAAFNATV 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 IVRFSTIHNSNESFDPATQEPG-VYVHCHILHEDHDMRPF-AMVL 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 5
AC0414
probable exported protein YP03409 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AC0414
R:Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibbail, R.W.; Holden, M.T.G.; Prentice, M.
  deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.
  11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
  Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AC0414
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-533 <KUP>
A:Cross-references: GB:AL590842; PIDN:CAC92639.1; PID:g15961336; GSPDB:GN00175
C:Genetics:
A:Gene: YP03409

```

```

Query Match 14.3%; Score 437.5; DB 2; Length 533;
Best Local Similarity 26.1%; Pred. No. 2.7e-24;
Matches 141; Conservative 70; Mismatches 209; Indels 121; Gaps 18;
QY 52 PLPIPPKQP-----RLTVNPVNGGEIWWYEVKIKPFTHQVYVVDLSADLVGYDGSPGP 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 PLPIPLLPDPDANGKTNL-NIQTGSYVWL-----PST-----ATQTWGNGALLDP 77
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 108 TPQVPRGVETVAFINNAAPNSVHLHGSFSRAAFGMAEDITEPGSKFYDYYPNRSAR 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 AIRLQGRKAVTIDITNALPBEATTVHNGLEIPGEVNDGCPQALIQPAKQVFAVBOQA 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 168 TIMYHNDHAMHITAENAYRQAGLYMLTDPAEDALNPSGGEEDIPMLTNSKQYTAGN 227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 TCWFPHHTSKTGHOVAMGIGLVLLIDSDSEFLPLPKQMGVDIVILLQDK-----L 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 228 VTTNGELN-----SFWGDVIVHNGQWPFKRVNPERKY-RFRFLDAVSRSGLYE 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 LDKHGVDVQDQVMTAAVGMFGDRMLTNGVPR-QDITRGWRVLRLLNGCAKRSNLN 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 277 ADTDAIDTRLPRKVIASDSGLLEHPADTSLITYISMAERYEVDFSDYAGKTIELNLG 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 ----ALSQGRPMYVIASDSGLLAEPVYVRELPILMGERREVLDVTRD--GQSLDLVTLPV 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

[illegible]

```

RESULT      6
AC3582
Probable blue-copper protein yack precursor [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AC3582
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Yanova,
  : Mazur, M.; Goldsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Leless
  : Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AC3582
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-494 >KUD>
A:Cross-references: GB:AO08918; PTDN:AL53822.1; PTD:q17984756; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110580
A:Map position: II

```

	Query Match	13.1%	Score 403;	DB 2;	Length 494;	
	Best Local Similarity	26.4%;	Pred. No. 8.3e-22;			
	Matches 137;	Conservative 71;	Mismatches 214;	Indels 96;	Gaps 16;	
OY	52 PLPIRPVQAPRLTIVNPVNGGEIWEYEVEIKRFTHQVYPDLGSADLVGVDGMSPEPTROY	111	:	:	:	:
Db	10 PLPLPRLPEVPDAS-----GIWRKLVQGRHS-FAKGSTASAGINCINGYLPLVRL	58	:	:	:	:
OY	112 PGCVTVVRFINNA-EAPNSVHLGFSFRAAFDGAEDITEGSGKDYVYPRQSARTLM	170	:	:	:	:
Db	59 MSG-ESVTLVENADEETLHMHGLFWPSHLDGPHNVIAPGAWEKVAAYNQSPASNW	117	:	:	:	:
OY	171 YHDHAMHTTAENARYGGAGLVMLTPAEDALNLPSCGYCEFDLPILITSKOYTANENV--	228	: :	: :	: :	: :
Db	118 FHPHLHGHTARAHMHGICGLMIVRDGKAERGLPETTYGVDDLPLVLQDRR-VIEBDANYA	176	:	:	:	:
OY	229 -TTNGELSNFMGDVHVHVGQPWFKNVEPRKRYRFLLDAVASRSGLYFADTDIDTRLP	287	:: :	:: :	:: :	:: :
Db	177 PDIMNLHGFRCDMLIVNGALIAPEARVPAAMVKRLFLNGANNRNHIRPAD-----GRP	230	:	:	:	:
OY	288 FKVIASDSGLEHPADTSILYSMAERYEVVDFSDYAGKTIELNLGSGSIGICTPDY	347	:	:	:	:
Db	231 LLVIASDGGLISQPSISIGLTISPGEREVLVDFSN--GEANDLVITYGNGSGDDL-----	284	:	:	:	:
OY	348 DNTDKVMRFVVAADDTTGDPDTSVVPANLRDVPPSPPTINTPROEFR-----	392	:	:	:	:
Db	285 ---HLMRPTV-DPALEGRVAKPVPVSFDGPAAPDEKLVSQRSSFDEFDERMAENMKLMRKQ	339	:	:	:	:
OY	393 -----GRGTGPTWTNG-----VAFADV---ONR	412				

```

Db      340 PSSNHAAGDDMDHDEMGSMAGMDEHMGSSAADAGPALDALTSGVMTALADKPFDER 399
OY      413 LIANPVGTVEWELINANGWTHPIHILHYDEKVIKSTSGNNATVPIYESGLKDYVL 472
Db      400 IDVEAKLSWETWEL--TSREMAHPFHIGASFRLTM-----NGKRAPHOTGMKDALI 453
OY      473 -GRRETVVVAHYAPFPGVYMFCHNLTHEDIDMAAF 509
Db      454 DGAELIVHFDEAKRSHPFMFCHLLEHEDYGMMAOF 491

```

```

RESULT 7
C64735
    probable copper-binding protein yack - Escherichia coli
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 08-Oct-1999
C:Accession: C64735; S45200
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
  A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426517
A:Accession: C64735
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-516 <BLAT>
A:Cross-references: GB:AF000121; GB:U00096; NID:g1786306; PIDN:AAC73234.1; PID:g1786306
R:Experimental source: strain K-12, substrain MG1655
R:Fujita, N.
    submitted to the EMBL Data Library, January 1994
A:Reference number: S45181
A:Accession: S45200
A:Molecule type: DNA
A:Residues: 1-463, 'LIRAC', 470, 'IPLR' <FUJ>
A:Cross-references: EMBL:D26562; NID:g473770; PIDN:BA05579.1; PID:d1066123; PID:g473770
A:Experimental source: strain K-12, substrain W310
C:Genetics:
A:Gene: yack
C:Keywords: copper

```

Query	Match	Similarity	13.1%	Score	401	DB	2	Length	516
Best Local	Similarity	28.8%	Pred.	No.1.2e-21					
Matches	136	Conservative	58	Mismatches	192	Indels	86	Gaps	17
QY	99	GXDGMSPGFTFVPPVPGVETVVRFINNAEAPNSVHLHGFSRAAFQMAEDITEPGSFKDY	158						
DB	68	GYNGLNLGPAVKLQSGKAVTDYVQLQTFEETLHHHGLTEVPGEVGGPGQIGIRPGKRSV	127						
QY	159	YTPNNGSATLWYHNDHAMITTAENAYRQAGLYMLDPAEDALNIPSGGEEDIMILTS	218						
DB	128	TINVDQPAATCFMHPHQHKGTRQYAMGLAGLVLEEDBELKMLPKQGLIDVPIVYOD	187						
QY	219	KOYTANGNLVTNNGELN-----SFWGDVYIHNGCPWPFKNYPERKY-RFRFLDAVRS	271						
DB	188	KKFSADGQI---DYQLDVTATAVGWFGDILLNGALY-P-QHAAPRGMLRLRLNCNARS	243						
QY	272	FGLYPADDTDAIDRLRPFKYIASDGLLHPADTSLIYISMARREYVDFSDYAKTIEL	331						
DB	244	--LNFNTSD---NRPLYIASDGLLPEPVYSELPRVLMGRFEVELEVND--NKPRDL	295						
QY	332	RNLGGIGIGTDTVDYDNTDKVMR---EYVADDTQOPDT-SVPPANLRLVPPSPPTNTP	387						
DB	296	VLLPVSQMGMAL-APFDKHPYMRIQPLAISAGALPDLSSLP-----LPDLGLGLV	348						
QY	388	RQFR-----FGRTG-----PFW	399						
DB	349	RRLQLSLMDEMLDMGQMQLMERYGDOAMAGMDHSQWIMGIMGCGNNHNNHGKCFDFNHAN	408						
QY	400	TINGVAFAVDVORLLANLVGVTEVERWELINAGNGTTPRHILVLDKVCISPTSGNNARTV	459						
DB	409	KINGQAFD--MKNKPRFAAKGQYERIVISGVDDMMILHPRHIGTQFRILS-----ENGKRP	462						


```

Db      439 PSMARFP--EYFENSOTLEMARDRSGISGOLLVNPVP 470
RESULT  14
E85961
suppressor of ftsI [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: E85961
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayn-
iller, L.; Grobbeck, E.U.; Davis, N.W.; Lim, A.; Diallanita, E.; Potamoudis, K.; Apoda-
tate 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E85961
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-470 <STO>
A:Cross-references: GB:E005174; NID:g12517586; PIDN:AAG58153.1; GSPDB:GN00145; UMGPR:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: sufi

Query Match          10.4%; Score 318.5; DB 2; Length 470;
Best Local Similarity 25.9%; Pred. No. 1.2e-15;
Matches 133; Conservative 72; Mismatches 234; Indels 75; Gaps 21;

OY    19 AVQASPVETSPATGHLEKRYAQISPOLYMTVELPIRPVAKQRLTYTNPVNGGEIMYYE 78
       |||::|||::|
Db     15 ALACAAVPLKASAAQ-----QQQLPVPPLLESr-----RCQPL-FMT 51

OY    79 VEIKETHQVPPDLGSADLVGYDGMSPGTFQVPEGVVFRTINNAEPNSVHLHGSSFS 138

```

QY 139 RAEDGMAEDITEPGSEFKDYIYYPNRQSARTLWYHDHAMHTAENAYRGQAQLYLMDPAE 198
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 109 PGPLMGPARMMSPNADWAPLYLRONATLWYHANTFNPTAQOYVNYGLAGMWLVDEVS 168
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 199 DALNPSTGYGEFDIPMLITSKQYTANGNLVTTNGELNSFGKDVLYHVNGQPMPFEKNVEPRK 258
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 169 KSLPFPNTGYVDDEFVIYIQDKRLDNFGRPEXNEBPSGGFVDDTLTVNVGSPPYEVSRSKW 228
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 259 YRFRIIDAAVSRSCELYFADDAIDTRLPFVYIASDGLLEHPADTSLYSMAERYRVY 318
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 229 VRLRLINSSNRKRYQLQMSD-----GRPILHYISDGQFLAPAVSVKOLSLAPGRRRL 282
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 319 FDFSDYAKTKIELRLNGISSIGITDYPDNDMKMFRVVADD-----TTQPDTSVY 370
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 283 VDMSN--GDEVSI--TGCEAST-----VDIRGFEESSIIIVSYLVLTLP--TELL 329
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 371 PANLEDVP--SPT---TNTPROEFRGTPTWTINGAFADVONRLANYPVGVERME 426
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 330 PLYVDSLPMRLLPFIMAGSPIRSRDISLGDPPIGNGLW--DY-NRIDVTAAQGWEWN- 386
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 427 LINGNGMTPIHITHLVDFKYISRTSGNNARTVMPY--ESGLKDVMVL-GRETIVV--E 481
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 387 TVRADE--PQAFHEIGEVFOI-----RNVGAMPFEREDRMKDPVWDGVQELLVVRGQ 438
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 482 AHYAFPCGVYMFCHNLIHEDHMAAMAFNAVLP 515
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 439 PSMAHF--FYFNQTLEMADRGSIGQLLVNVP 470
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
RESULT 15
safti protein precursor - Escherichia coli
C:Species: Escherichia coli
C:date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 08-Oct-1999
C:accession: G65008; S20461
F:Blaetter, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: G65088
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-470 <BLAT>
A:Cross-references: GB:A6000384; GB:U00096; NID:92367186; PIDN:AAC76053.1; PID:91789394;
R:Experimental source: strain K-12, substrain MG1655
R:Coleman, J.
Mol. Gen. Genet. 232, 295-303, 1992
A:Title: Characterization of the *Escherichia coli* gene for 1-acyl-sn-glycerol-3-phosphat
A:Reference number: S20459; MUID:92212294
A:Accession: S20461
A:Molecule type: DNA
A:Residues: 1-89 <COL>
A:Cross-references: GB:M63491; NID:9147296; PIDN:AAA24398.1; PID:9147299
C:Genetics:
A:Gene: *sufl*

Query Match 10.3%; Score 315.5; DB 2; Length 470;

Best Local Similarity 25.9%; Pred. No. 2, 1e-15;
Matches 133; Conservative 71; Mismatches 235; Indels 75; Gaps 21;

```
OY 19 AVQASVPETSPATGHLERKVAQISFOYPMFTVPLPIPVKOPRLTVTPNVNGQETWYE 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 15 ALCAGAVPLKASAAQ-----00PLVPPLLES-----RCQPL-FWT 51

OY 79 VEIKPFTHQVYDLSADLVGDSGSPGTFQVPRGVEVTVREINNAEAPNSVHLGSFS 138
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 52 VQ---RAHMSFPGTRASWYGIKRYLGRKWDVKKLYSNRLTENVSMTVAILOV 108

OY 139 RAAPGMAEDITEPGSFKDYYPNROSARTLWYHDHAMHTAENAYRGOAGLYMLTDPAE 198
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 109 PGRLLMGGRPARMMSPNADMAPVLPIRQNAATLWYHANTPRTAQOYVNGLAGMMLVEDEVS 168

OY 199 DALNLPSCYGEFDIPHILTSKQYLTANGELNSFWGVYIHVNGOPWPKNVPRK 258
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 169 KSLPIPNHGVDDFVYIIOCKRLDNFGTEPEYNEPGSGGFVDTLLVNGVQSPYEVSRCQ 228

OY 259 YRFRFLDAVSRSGFYFADTDALDTRLPFKVIASDSGLLEHPADTSLLYISMAERYEV 318
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 229 VRLRLINASNSRYLOLQND-----GRPLHVISGDQGLPAPVSVKOLSLAGERRETI 282

OY 319 PDESDYAGKTIELRLNGSIGIGITDIDYDNTDKVRFVAVD-----TTQPDTSVY 370
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 283 VQMSN--GDEVSI--TCGEASI-----VDRIKGFEPSPSILVSTLVLTLP--TGLL 329

OY 371 PANLKDVPFP-SPT---TTPROFRFGRTGPTWTINGVAFADYQNRLLANVPVGTIERME 426
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 330 PLVTDSLPRMLLPTETIMASPIRSRDISLGDDPGINGQLM-DV-NRIDVTAQGTWERM- 386

OY 427 LINAGWTHPIHILVDKVISRTSGNNARTWMPY--ESGLKDVWL--GRRETVV--E 481
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 387 TVRADE--PQAHIEGVMTQI-----RVNVCAMPPEPDGKMDTYWVDGQVELLYVFGQ 438

OY 482 AHYAPFPGVYMFCHNLHEDHDMAAFNATVLP 515
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 439 PSMHFP--FYFNSQTFLEMADRSGISQOLLVNPVP 470
```

Search completed: October 3, 2002, 10:55:35
Job time: 79 sec

PT Bilirubin oxidase prep. useful as an analytical enzyme - by
PT culturing bilirubin oxidase in transformant culture

PS Claim 1; Page 29-32; 32pp; Japanese.

XX The sequence encodes bilirubin oxidase. The protein produced has a
CC 38 amino acid signal peptide which is removed to give the mature
CC protein (AAR40843).

XX Sequence 572 AA;

Query Match 100.0%; Score 3067; DB 14; Length 572;
Best Local Similarity 100.0%; Pred. No. 1.2e-270;
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MEKHTLGAALSLFNSNAVQASVPETSPATGHLFKRAQISPOVMPTVLPPIPVKQ 60
DB 1 mkkhclgaalsllfnsnagaspvetspacghlfrvaqispqpmfvtvlpipvq 60
OY 61 PRLTVTNPVNGOEIWEYVEIKFTHQVYVPLDGSADLVGDMSPQPTQVGRVETVVR 120
DB 61 prltvtnpngoeiwyveikfthqvypdlgsadlvvgdmspqptqygrvetyvr 120
OY 121 FNNNAEAPSVVHLHGSFSAAFDGMADTEPGSFQDYTPYRNQSAKTLWYHDAMHTA 180
DB 121 fnnnaeapsvvhlhgsfsaafdgmaedtepgsfqdytprnqsaktlwyhdamhta 180
OY 121 fnnnaeapsvvhlhgsfsaafdgmaedtepgsfqdytprnqsaktlwyhdamhta 180
OY 181 ENAYGQAGLWLTDPADALNLPSCGCEFDIPMILTSQYANGNLVTNGELNSFMD 240
DB 181 enaygqaglwltdpadalnlpscgcefdipmiltsqyngnlvtngelnsfmd 240
OY 241 VTHVNGQPFKFNVEPRKRFREFLDAAVRSFGLYFADTDAIDTRLPEFVIASDGLLEH 300
DB 241 vthvngqpfpkfnveprkrfrefldaaavrsfglyfadtdaidtrlpfviassdglleh 300
OY 301 PADTSLIYSMAERYEVFDESDYAGKTIELNLGSGICIGTDYDNTDKVMRFVAD 360
DB 301 padtsllysmaeryevdfdsyagktielnlgsigicigtddydyntdkvmrfvad 360
OY 361 DDTQDTSVPANLDRVPPPTNTNPROFRGRTGPTWINGVAFADVQNLNANVP 420
DB 361 dttdqdtsvpanldrvppptntnprofrgrtgptwingvafadvqnlannvp 420
OY 421 TVERMELINAGMWTPIHILVDFKVISRTSGNNAKRTVMPEYSGLKDVWLGRTVVV 480
DB 421 tvermelinagmwtphihlvdfkvisrtsgnnarkrtvmpeysgldkvwlgretvvv 480
OY 481 EAHYAPFPGVYVFRCHNLHEDHDMMAAFNATVLPDYGNATVFPVDMELMQARPELG 540
DB 481 eahyapfpgvymfrchnlhedhdmmaafnatvlpdygnatvfpvdmelmqarpeylg 540
OY 541 EFOAGSGFVQAVTERIOTMAEYRYAADE 572
DB 541 efqagsgfvavteriotmaeyryaaade 572

```

RESULT 2

AAB81506
ID AAB81506 standard; Protein; 578 AA.

AC AAB81506;

DT 18-JUN-2001 (first entry)

DE Bilirubin oxidase.

KW Bilirubin oxidase; phenol oxidizing enzyme; phenol oxidase B;

KW detergent; paper production; pulp production; textile; food industry;

OS Unidentified.

PN WO200121748-A1.

XX 29-MAR-2001.

XX 06-SEP-2000; 2000WO-EP08840.

XX 22-SEP-1999; 99EP-0203120.

XX (UNIL) UNILEVER NV.

XX (UNIL) UNILEVER PLC.

XX (HIND-) HINDUSTAN LEVER LTD.

XX Convents D, Doornik M, De Vries CH, Wang H;

XX WPI; 2001-273462/28.

PT New detergent compositions comprising a phenol oxidizing enzyme useful
PT in detergent or cleaning compositions, fiber treatment, processing,
PT finishing or production, paper and pulp production, or in starch
PT processing applications.

XX Example 4; Fig 4; 46pp; English.

CC The present sequence was used for comparison with the stachybotrys
CC charitatum phenol oxidizing enzyme. The invention relates to detergent
CC compositions comprising one or more surfactants and a phenol oxidizing
CC enzyme having at least 68% identity to the stachybotrys charitatum phenol
CC oxidizing enzyme. Phenol oxidizing enzymes may be used in the detergent,
CC paper, pulp, textile and food industries. They are used for preventing
CC the transfer of dyes in solution from one textile to another during
CC detergent washing, or in modifying the colour associated with dyes and
CC colourant compounds having different chemical structures, such as in pulp
CC and paper bleaching, bleaching the colour of stains on fabric and in
CC detergent and textile applications.

XX Sequence 578 AA;

Query Match 65.5%; Score 2010; DB 22; Length 578;
Best Local Similarity 65.4%; Pred. No. 2.8e-174;
Matches 383; Conservative 71; Mismatches 110; Indels 22; Gaps 10;

```

OY 1 MEKHTLGAALSLFNSNAVQASVP--ETSPATGHLFKRAQISPOVMPTVLPPIPV 57
DB 1 mskqalgaalglavlgssvdaravgrstlmpsg-ltkrtqlspalalyevldpipp 59
OY 58 VNOPRLTVNPNVNGOEIWEYVEIKFTHQVYVPLDGSADLVGDMSPQPTQVGRVET 117
DB 58 vnoprltvnppngoeiwyveikfthqvypdlgsadlvvgdmspqptqygrvety 117
OY 60 KAPN-typnptgedillyemeitpfshqlypdlpanamvydgmspgptliypgtes 118
DB 60 kapn-typnptgedillyemeitpfshqlypdlpanamvydgmspgptliypgtes 118
OY 118 VYRFINNAE--APNSVHLHGSFSAAFDGMADITEPGSFQDYTPYRNQSAKTLWYHDA 175
DB 118 vyrfinsngtspnsvhlhgsfsaafdgmaeditepgsfqdytprnqsaktlwyhda 178
OY 176 MHTTAENATRGAGLWLTDPADALNLPSCGCEFDIPMILTSQYANGNLVTNGELN 235
DB 176 mhttaenaymgagayymldqpaedalnlpssygeidpvlalakrynaogtlfinsngs 238
OY 236 SFGGVYHNGQPFKFNVEPRKRFREFLDAAVRSFGLYFADTDAIDTRLPEFVIASD 295
DB 236 sfvgdvihngqpfpkfnveprkrfrefldaaavrsfglyfadtdaidtrlpfviass 298
OY 299 gllghpADTSLIYSMAERYEVFDESDYAGKTIELNLGSGICIGTDYDNTDKVMR 355
DB 299 gllghpvdtdtlyismaerewevldstagsidlrlnpgs-dglgyepelndtkvmr 357
OY 356 FVADDTQPTDTSVPANLDRVPP-----SPTNTNPROFRGRTGPTWINGVAF 407
DB 356 fvdvdelesdpsevpnanldrvfpdggnwdbandld--elftgrangvgtlmgvlf 415
OY 408 DVONRLANVPYGTVERWELINAGMWTPIHILVDFKVISRTSGNNAKRTVMPE-SGL 466
DB 416 dvenrlrnvpdtevlwrlennsgwchpnhlvdfkvisrst--argvpeyaagl 472

```


[illegible]

RESULT 6

ID AA95538 standard; Protein; 627 AA.
 AC AA95538;
 DT 10-OCT-2000 (first entry)
 DE Bipolaris spicifera phenol oxidising enzyme.
 KW Phenol oxidising enzyme; detergent; bleaching.
 OS Bipolaris spicifera.
 PN MO200039306-A2.
 XX 06-JUL-2000.
 XX 20-DEC-1999; 99WO-EPI0287.
 XX PF 23-DEC-1998; 98US-0220871.
 PR 23-JUN-1999; 99US-0338723.
 XX (UNIL) UNILEVER NV.
 PA (UNIL) UNILEVER PLC.
 PI (HIND-) HINDUSTAN LEVER LTD.
 FI Bodie EA, Van Der Velden S, De Vries CH, Wang H;
 XX MPI; 2000-514528/46.
 DR N-PSDB; AAA50020.
 XX
 PT Detergent composition comprising novel phenol oxidising enzyme obtained
 PT from fungus or bacteria, useful for pulp and paper bleaching, bleaching
 XX color of stains on fabric and for anti-dye redeposition
 XX
 PS Claim 8; Fig 3; 45pp; English.
 XX
 CC The present sequence is that of the *Bipolaris spicifera*
 CC phenol oxidising enzyme. The invention relates to detergent
 CC compositions comprising novel phenol oxidising enzymes that have at
 CC least 60% identity with the phenol oxidising enzyme of *Stachybotrys*
 CC *chartarum* (see AA95537), and which are obtained from a bacterium,
 CC yeast or non-*Stachybotrys* fungus, especially *B. spicifera*,
 CC *Curvularia pallens* (see AA95539) and *Amerosporium atrum* (see
 CC AA95540). The phenol oxidising enzyme is capable of modifying the
 CC colour associated with dyes or coloured compounds, and can be used
 CC for pulp and paper bleaching, for bleaching the colour of stains on
 CC fabric and for anti-dye transfer in detergent and textile
 CC applications. It may also be capable of modifying the colour in
 CC the absence or presence of an enhancer. Expression vectors and
 CC host cells comprising a nucleic acid encoding a phenol oxidising
 CC enzyme, methods for producing the phenol oxidising enzyme, and
 CC methods for constructing expression hosts are provided.
 XX
 XX Sequence 627 AA:

Oy	43	SPOYR-ETVLEPILPRVXQOPRLTYTNPNVNGEIMUYEVEKLKPRFHYYPRDLSADLYGD	101
Dd	63	speyplilrqpilrpbakepn-kltnpytkelwyeyelvkprfqqyyslparllygyd	121
Oy	102	GMSPEPTQVPRGYEVYVRFINNAEAPNSVHLHGFSRAAFEDGMAEDITEGSEFKDYYP	161
Dd	122	gispeplliivrptgeaavryfingddresslilbhspsrarpfdgwadmmkgyekdyyp	181
Oy	162	NROSARLTWYNDHAMHTLEANAAYGOGGLYMLTPRADALNLPESGVEPRLPMLTNSKOY	221
Dd	182	mgaerflwyndhamhvaenaeylrgaqayliltbpaedaqlrpsgykylpdlvlskyy	241
Oy	222	TANGNLVTTNSELNSFEMGDVLIHVNGOPMPEKNEPRKYRFRFELDAAVSFGELVPADTDA	281
Dd	242	nadgkltsvsgedksvwgdlilhvngqprflnvbrgryrlrflnaasrnfaifyvkqn	301
Oy	282	IDTRLPEFVINSDSGLLENHPADTSLATYSMAEREVEYFEDSDVAGKTELRLNLSGSGI	341
Dd	302	tatrlpqlvlsdaeglllthnpqtsdmyaaeyelvdfdarayagqlldltnf-akangi	360
Oy	342	GTDTDYDUTDKVVMRFVVAADDTTQDTSVVPANLNDVFEPSPTTNPQRFGRGTGPTWI	401
Dd	361	gtdddyantdkvmfthvssqcv-vdnusvpeqldsqifpdktdidhhfrfntngswl	419
Oy	402	NGVAFADVQNMLLANVPRGTYERBELINAGGWNPHNHLNYPKVLSPKSGNNARVYR	461
Dd	420	nglgladavenlrvlkvprglvlelwnelssgwsrhlrvhvlvdrvaaylvgdegltywmp	479
Oy	462	YE-SGLKDVWMLGRRETVWEAHNAPRGVYMFCHNLIHEDHDMMAFAFNATVLPDDGN	520
Dd	480	yeaaglkdvwmlgshetvleahyapwdgywmfchnlidhgdmmaafvtrklqntgyn	539
Oy	521	ATV-FVDMELMARPVELGEFPAQSGQSVAVTERIOTMAETRYA	568
Dd	560	etldhdpedprwaiprltaglarslilseeslarrvelaaleqps	588

RESULT 7

TD	AAV96762 standard; Protein: 627 AA.
XX	
AC	AAV96762;
XX	
DT	09-OCT-2000 (first entry)
XX	
DE	Bipolaris spicifera phenol oxidizing enzyme.
XX	
KW	Phenol oxidizing enzyme; colour; dye; modification; detergent; strain:
KW	pulp; paper bleaching.
XX	
OS	Bipolaris spicifera.
XX	
PN	MO200037654-A2.
XX	
PD	29-JUN-2000.
XX	
PF	20-DEC-1999; 99WO-US31009.
XX	
PR	23-DEC-1998; 98US-0220871.
PR	23-JUN-1999; 99US-0338723.
XX	
PA	(GEMV) GENENCOR INT INC.
XX	
PI	Wang H, Bodie EA;
XX	
DR	WPI; 2000-452191/39.
DR	N-PSDB; AAA51315.
XX	
PT	New phenol oxidizing enzyme for modifying colors associated with dyes
PT	or colored compounds, is obtained from fungus and is encoded by a
PT	nucleic acid comprising a specific nucleotide sequence
XX	

PS Claim 8; Fig 3; 45pp; English.

CC This sequence is the *Bipolaris spicifera* phenol oxidizing enzyme.
CC Phenol oxidizing enzymes encoded by nucleic acid sequences which
CC hybridize to the *Stachybotrys chartarum* DNA (AA9513) are claimed, as
CC long as the enzyme is capable of modifying the colour associated with
CC dyes or coloured compounds. The enzymes are useful in detergent
CC compositions and for modifying colors associated with dyes or coloured
CC compounds which occur in stains in a sample. The enzymes are also useful
CC for pulp and paper bleaching, anti-dye transfer in detergent and other
CC textile applications.

CC Sequence 627 AA:

Query Match 58.7%; Score 1799; DB 21; Length 627;
Best Local Similarity 62.9%; Pred. No. 5.5e-155;
Matches 333; Conservative 67; Mismatches 123; Indels 6; Gaps 6;

QY 43 SPQYPM-FTVPPPIPVKQRLVTNPVNGOEIWEYVEIKFPTHOVYDLSADLVGYD 101
DB 63 SPEYPIIFRQPIIPKAPKPN-KLNPVFNKELWYVEIKFPTGVPSLPATLVGYD 121
QY 102 GMSPGPTFOVPRGVEYVRFINNAEAPNSVHLGFSRAAFDGMADITEPGSFKDYYP 161
DB 122 GISPGLTIVPGTEAVRVFNGDRESSIHHSRSPRAPFDGWADMLMKGEYKDYYP 181
QY 162 NQOSARTLWYHHAHMTENAYRGQAGLYMLTPAEDALNLPSCVGEEDIMILTSKY 221
DB 182 NQGAFTLWYHHAHMTENAYRGQAGLYMLTPAEDALNLPSCVGEEDIMILTSKY 241
QY 222 TANGNLVTNGELNSFWGVIVHNGOPMPKFNVEPRKRYFRFLDAVSRFSGLYFADTDA 281
DB 242 NADGGLKTSVGEKSWGDIHNVGQWPFILNVEPKYLRFLNAVSFNALYFKGQN 301
QY 282 IDTRLPEFKVIASDGLLEHAPADTSLYISMAERYEVVFPFSDYACTIELRLNGSIGI 341
DB 302 TATRLPIFYIASDGLLTPVQTSIDMYAAERYELVIFAPYAGQLIDRLNF-AKANGI 360
QY 342 GTDTYDNTDKVMRFVADDTQPTSVVPANLRDVPSPPTNPFROFRGRTGPTWTI 401
DB 361 GTDDYANTDKVMRFVADDTQPTSVVPANLRDVPSPPTNPFROFRGRTGPTWTI 419
QY 402 NGVAFADVONRLANVPVGTVERMELINAGCWTHPIHILVDFEKYISRTSGNNATVWP 461
DB 420 NGIGTADVONRLANVPVGTVERMELINAGCWTHPIHILVDFEKYISRTSGNNATVWP 479
QY 462 YE-SGLKOVVWLGRETVVVEAHYAPFGVYMFHCHNLHEDHMAATNATVLPDYGIN 520
DB 480 YEAGLKVWVWLGRETVVVEAHYAPFGVYMFHCHNLHEDHMAATNATVLPDYGIN 539
QY 521 ATV-FVDPVEELMOARPYELGEFOAGSOFVQVATERIQTMAEYRPA 568
DB 540 ECTDHDPEDPWRSARPIAGLTARSGIFSEESTARVNEALAEQPS 588

RESULT 8

AA95539
ID AA95539 standard; Protein; 627 AA.

AC AA95539;

DT 10-OCT-2000 (first entry)

XX Curvularia pallescens phenol oxidising enzyme.

XX Phenol oxidising enzyme; detergent; bleaching.

XX Curvularia pallescens.

OS WO200039306-A2.

XX 06-JUL-2000.

XX 20-DEC-1999; 99WO-EP10287.

XX 23-DEC-1998; 98US-0220871.

XX 23-JUN-1999; 99US-0338723.

XX (UNITL) UNILEVER NV.

XX (UNITL) UNILEVER PLC.

XX (HIND-) HINDUSTAN LEYER LTD.

XX Bodie EA, Van Der Velden S, De Vries CH, Wang H;

XX WPI; 2000-514528/46.

XX N-PSDB; AAA50021.

XX Detergent composition comprising novel phenol oxidising enzyme obtained

XX from fungus or bacteria, useful for pulp and paper bleaching, bleaching

XX color of stains on fabric and for anti-dye redeposition

XX Claim 8; Fig 10; 45pp; English.

CC The present sequence is that of the *Curvularia pallescens*
CC phenol oxidising enzyme. The invention relates to detergent
CC compositions comprising novel phenol oxidising enzymes that have at
CC least 60% identity with the phenol oxidising enzyme of *Stachybotrys*
CC *chartarum* (see AA95537), and which are obtained from a bacterium,
CC yeast or non-*Stachybotrys* fungus, especially *C. pallescens*,
CC *Bipolaris spicifera* (see AA95538) and *Amerosporium atrum* (see
CC AA95540). The phenol oxidising enzyme is capable of modifying the
CC colour associated with dyes or coloured compounds, and can be used
CC for pulp and paper bleaching, for bleaching the colour of stains on
CC fabric and for anti-dye transfer in detergent and textile
CC applications. It may also be capable of modifying the colour in
CC the absence or presence of an enhancer. Expression vectors and
CC host cells comprising a nucleic acid encoding a phenol oxidising
CC enzyme, methods for producing the phenol oxidising enzyme, and
CC methods for constructing expression hosts are provided.

CC Sequence 627 AA:

Query Match 58.5%; Score 1795; DB 21; Length 627;
Best Local Similarity 63.1%; Pred. No. 1.3e-154;
Matches 334; Conservative 63; Mismatches 126; Indels 6; Gaps 6;

QY 43 SPQYPM-FTVPPPIPVKQRLVTNPVNGOEIWEYVEIKFPTHOVYDLSADLVGYD 101
DB 63 SPEYPIIFRQPIIPKAPKPN-KLNPVFNKELWYVEIKFPTGVPSLPATLVGYD 121
QY 102 GMSPGPTFOVPRGVEYVRFINNAEAPNSVHLGFSRAAFDGMADITEPGSFKDYYP 161
DB 122 GISPGLTIVPGTEAVRVFNGDRESSIHHSRSPRAPFDGWADMLMKGEYKDYYP 181
QY 162 NQOSARTLWYHHAHMTENAYRGQAGLYMLTPAEDALNLPSCVGEEDIMILTSKY 221
DB 182 NQGAFTLWYHHAHMTENAYRGQAGLYMLTPAEDALNLPSCVGEEDIMILTSKY 241
QY 222 TANGNLVTNGELNSFWGVIVHNGOPMPKFNVEPRKRYFRFLDAVSRFSGLYFADTDA 281
DB 242 NADGGLKTSVGEKSWGDIHNVGQWPFILNVEPKYLRFLNAVSFNALYFKGQN 301
QY 282 IDTRLPEFKVIASDGLLEHAPADTSLYISMAERYEVVFPFSDYACTIELRLNGSIGI 341
DB 302 TATRLPIFYIASDGLLTPVQTSIDMYAAERYELVIFAPYAGQLIDRLNF-AKANGI 360
QY 342 GTDTYDNTDKVMRFVADDTQPTSVVPANLRDVPSPPTNPFROFRGRTGPTWTI 401
DB 361 GTDDYANTDKVMRFVADDTQPTSVVPANLRDVPSPPTNPFROFRGRTGPTWTI 419
QY 402 NGVAFADVONRLANVPVGTVERMELINAGCWTHPIHILVDFEKYISRTSGNNATVWP 461
DB 420 NGIGTADVONRLANVPVGTVERMELINAGCWTHPIHILVDFEKYISRTSGNNATVWP 479

XX 06-JUL-2000.
PD
XX 20-DEC-1999; 99WO-EPI0287.
PF
XX 23-DEC-1998; 98US-0220871.
PR 23-JUN-1999; 99US-0338723.
XX
PA (UNITL) UNILEVER NV.
PA (UNITL) UNILEVER PLC.
PA (HIND-) HINDUSTAN LEVER LTD.
PI Bodie EA, Van Der Velden S, De Vries CH, Wang H;
XX
XX WPI: 2000-514528/46.
DR N-PSDB; AAs50018, AAs50019.
XX
PT Detergent composition comprising novel phenol oxidizing enzyme obtained
PT from fungus or bacteria, useful for pulp and paper bleaching, bleaching
PT color of stains on fabric and for anti-dye redeposition
XX
XX Claim 2: Fig 5A-B; 45bp; English.

CC The present sequence is that of a phenol oxidizing enzyme of
CC Stachybotrys chartarum MUC1.38898. A claimed detergent composition
CC comprises a phenol oxidizing enzyme having at least 60% identity to
CC the present sequence, and preferably obtained from a bacterium,
CC yeast or non-stachybotrys fungus, especially Bipolaris spicifera,
CC Curvularia pallescens or Amersporium atrum (see AAY95538-40). The
CC phenol oxidizing enzyme is capable of modifying the colour
CC associated with dyes or coloured compounds, and can be used for
CC pulp and paper bleaching, for bleaching the colour of stains on
CC fabric and for anti-dye transfer in detergent and textile
CC applications. It may also be capable of modifying the colour in
CC the absence or presence of an enhancer. Expression vectors and
CC host cells comprising a nucleic acid encoding a phenol oxidizing
CC enzyme, methods for producing the phenol oxidizing enzyme, and
CC methods for constructing expression hosts are provided.

SO Sequence 594 AA;

Query Match 56.2%; Score 1724.5; DB 21; Length 594;
Best Local Similarity 60.3%; Pred. No. 3,1e-148;
Matches 322; Conservative 75; Mismatches 128; Indels 9; Gaps 7;

OY 43 SPOYP-MFTVPLPIPVKOPRLVTNPVNGQELWYEVETKPTTHOYVYDPLGSADLVGYD 101
DB 57 SPPYNLLYRNALPIPVKQPKMLITNPVNGQELWYEVETKPTTHOYVYDPLGSADLVGYD 116
OY 102 GMSPGPTFOVPRGVEVTFVREINNAEAPNSVHLHGSFSAFDFGMAWDITEPGSFKDYYP 161
DB 117 GMSPGPTFOVPRGVEVTFVREINNAEAPNSVHLHGSFSAFDFGMAWDITEPGSFKDYYP 176
OY 162 NROSARTLWYHDAHMTAENAYRGAGLYMLTDPADALNLPSSGGEEDIPMILTSKOY 221
DB 177 NYGSARLLWYHDAHMTAENAYRGAGLYMLTDPADALNLPSSGGEEDIPMILTSKOY 236
OY 222 TANGNLTNGELNSFWGVHINAGPMPKRNPRKRYRFRFADAVSFGIYFDTPA 281
DB 237 NADGTLTSTEGEDQDLWGVHINAGPMPKRNPRKRYRFRFADAVSFGIYFDTPA 296
OY 282 IDTRLEFKIASDSGLLEPRADTSLYISMAERYEVVDFSDAGRTIELNLGSGIGC 341
DB 297 PNVRILPFGYLAASDAGLLQPVGVSOLYLAEAEEYELIDFTNAGTDLDIRV-AETND 355
OY 342 GIDTVDNMDKVRVAVADDTTQPTGSVPANLRDVPFSPSTTNP--ROFRFGRTGPTW 399
DB 356 GDEDEYARLEVMRIVSSGTYE-dnsqybstlrdvpfp-phkegpadckfkfetersngyh 413
OY 400 TINGVAFADYONRLANVPVGTWERELINAGGWHPIHIIHVDKVISRTSGNNARY 459
DB 414 LINDVGLADVNERVLAKPELIVTEWELNSGGWSHPVHILVDKILKTRIGRG 471

OY 460 MPYES-GLKDVWLAGRETVVEAHYAPFEGVYMFCHNLHEDHDMMAFNATVLPDYG 518
DB 472 MPYESAGLKDVWLAGRETVVEAHYAPFEGVYMFCHNLHEDHDMMAFNATVLPDYG 531
OY 519 YNATVFPDMEELMWARPELGEFOASGOFVQAVATERIQTMAERYPPAAADE 572
DB 532 YLQEDFEDPMNPWRVAVPYRNDFNARGNFSASSTARVQELAEQEPYRIDE 585

RESULT 13

AAY96761
ID AAY96761 standard; Protein; 594 AA.
XX

AC AAY96761;
XX

DT 09-OCT-2000 (first entry)
XX

DE Stachybotrys chartarum phenol oxidizing enzyme.
XX

KW Phenol oxidizing enzyme; colour; dye; modification; detergent; stain;
KW pulp; paper bleaching.
XX

OS Stachybotrys chartarum.
XX

PN WO200037654-A2.
XX

PD 29-JUN-2000.
XX

PF 20-DEC-1999; 99WO-US31009.
XX

PR 23-DEC-1998; 98US-0220871.
PR 23-JUN-1999; 99US-0338723.
XX

PA (GENV) GENENECOR INT INC.
XX

PI Wang H, Bodie EA;
XX

DR WPI: 2000-452191/39.
DR N-PSDB; AAs51313, AAs51314.
XX

PT New phenol oxidizing enzyme for modifying colors associated with dyes
PT or colored compounds, is obtained from fungus and is encoded by a
PT nucleic acid comprising a specific nucleotide sequence
XX
PS Claim 1: Fig 5A-B; 45bp; English.

CC This is the sequence of Stachybotrys chartarum phenol oxidizing enzyme.
CC Phenol oxidizing enzymes encoded by nucleic acid sequences which
CC hybridize to the coding DNA are claimed, as long as the enzyme is capable
CC of modifying the colour associated with dyes or coloured compounds. The
CC enzymes are useful in detergent compositions and for modifying colors
CC associated with dyes or colored compounds which occur in stains in a
CC sample. The enzymes are also useful for pulp and paper bleaching,
CC anti-dye transfer in detergent and other textile applications.
XX
SO Sequence 594 AA;

Query Match 56.2%; Score 1724.5; DB 21; Length 594;
Best Local Similarity 60.3%; Pred. No. 3,1e-148;
Matches 322; Conservative 75; Mismatches 128; Indels 9; Gaps 7;

OY 43 SPOYP-MFTVPLPIPVKOPRLVTNPVNGQELWYEVETKPTTHOYVYDPLGSADLVGYD 101
DB 57 SPPYNLLYRNALPIPVKQPKMLITNPVNGQELWYEVETKPTTHOYVYDPLGSADLVGYD 116
OY 102 GMSPGPTFOVPRGVEVTFVREINNAEAPNSVHLHGSFSAFDFGMAWDITEPGSFKDYYP 161
DB 117 GMSPGPTFOVPRGVEVTFVREINNAEAPNSVHLHGSFSAFDFGMAWDITEPGSFKDYYP 176
OY 162 NROSARTLWYHDAHMTAENAYRGAGLYMLTDPADALNLPSSGGEEDIPMILTSKOY 221
DB 177 NYGSARLLWYHDAHMTAENAYRGAGLYMLTDPADALNLPSSGGEEDIPMILTSKOY 236

Db 177 nysarllwyhahfntaenayfygagaylndaeedlglpsygefdipllllaky 236
QY 222 TANGNLVTNGELNSWGDVIVHNGQMPFKVNEPRKRYFRFLDAVSRSGLYFADTDA 281
Db 237 naddgltrstegedqdlwgdvlnvngpwpflnvqpkkyfrfinaavstrawlllyrtas 296
QY 282 IOTRLPFKVIASDGLIEHPADTSLIYISMARERYEVPFSDYAKTIELRLGSGISGI 341
Db 297 pvrilpfigvlaadagllqgarpvtslnlylaaeryellldftfhaqgclldlrnv-aetndv 355
QY 342 GNDTODNDMDKWRFPVADDTOPDTSVVPANRDVPFSPPTNTP--KQFRGRTPTM 399
Db 356 gdedeafartlemrftvssggtve-dnsqypstlrtdyfp-pkhegpadkhfkertsnghy 413
QY 400 TINGVAFAVDONRLANVPGVTERMELINAGNGWTHPIHILVDEKVISRPSGNARTV 459
Db 414 lndvgfadvnervelakpdlgcvewelenssgwshpnhlhfklkrtggrg--qv 471
QY 460 MPYES-GLKDVVWLGRRFVVVEAHTAPPQVYMFCHNLIEDHDMMAFNATVLPDYG 518
Db 472 mpyesaglkdvwylgrgetlleahygpwtgaymwhchllhedndmavfntameekg 531
QY 519 YNATVPVDMELMOWARPELGEFOAGSOFVSQAVTERIQTMAEVRPYAADE 572
Db 532 ylgdedfpmnpkwaavpynrnfharagnfsesitarvqelaegepyrnde 585

RESULT 14

AAU04831
ID AAU04831 standard. Protein: 474 AA.

XX AAU04831;

DT 26-SEP-2001 (first entry)

XX Micromonospora everninnomycin biosynthetic enzyme evrg.

XX Everninnomycin; antibiotic; bottle-neck gene; orthomycin;

KW fermentation; tailoring gene product; oxidase;

KM evrg.

XX Micromonospora carbonacea var. africana.

OS WO200151639-A2.

PN 19-JUL-2001.

XX 12-JAN-2001: 2001WO-US01187.

XX 12-JAN-2000: 2000US-0175751.

XX (SCHE) SCHERING CORP.

XX Hosted TJ, Horan AC, Wang TX;

XX WPI: 2001-442147/47.

XX N-PSDB: AAS08693.

XX New nucleic acid molecules encoding everninnomycin pathway gene

XX PT products, useful for improving yields of everninnomycin, to produce new

XX PT everninnomycin and as probes to identify homologous sequences

XX Claim 19; Fig 11; 109pp; English.

XX The sequence represents a Tailoring gene product, an oxidase,
CC evrg. The protein comprises one of 98 enzymes of the
CC everninnomycin antibiotic biosynthetic pathway. A vector comprising a
CC M. carbonacea everninnomycin biosynthetic pathway resistance gene product
CC is useful for selecting for a transfected or transformed host cell. An
CC integrative version of the vector is useful for introducing a
CC everninnomycin pathway gene (a bottle-neck gene) into an actinomycete of
CC the genus Micromonospora. The DNA encoding the biosynthetic proteins is
XX useful for synthesizing novel everninnomycin-related compounds, arising

CC from modifications of the DNA sequence designed to change glycosyl and
CC modified iscellithic acid groups contained in everninnomycin, for
CC expressing functional or mutant everninnomycin biosynthetic enzyme for
CC evaluation, diagnosis and preferably biosynthesis of everninnomycin or
CC other secondary metabolic products, improving the yield of everninnomycin
CC and to produce novel everninnomycin and also as a hybridisation probe to
CC identify homologous sequences. The encoded polypeptides are useful for
CC combinatorial biosynthesis to generate libraries of orthomycins, e.g.
CC everninnomycin analogues/homologues and drug discovery. The
CC DNA encoding the integrase allows for increasing a given gene dosage. The
CC integrative vector can be used to permanently integrate copies of a
CC heterologous gene of choice into chromosomes of different hosts and to
CC integrate genes which increase the yield of known products or to generate
CC novel products such as hybrid antibiotics or other novel secondary
CC metabolites. The vector can also be used to integrate antibiotic
CC resistance genes in order to carry out bioconversions with compounds to
CC which the strain is normally sensitive and is thus useful in fermentation
CC processes involving e.g. Streptomyces antibiotics.

XX Sequence 474 AA:

Query Match 21.8%; Score 669.5; DB 22; Length 474;
Best Local Similarity 34.1%; Pred No. 3.3e-52;
Matches 175; Conservative 76; Mismatches 197; Indels 65; Gaps 18;

QY 12 SLLENSNAVQASPVPETSPATGH--LFRKVAQISPOYEMFTVPLPIPPVKOPRLTVTNP 68

Db 4 sltrdagaqaaspaspapnaghaaprsstltvavrftepmvp-----plrlpvr 59

QY 69 VNGQELWYEEVEIKPFTTHOYVPDLGSADLVGDGMSPGTFPOVPRQVEVVFIRINAEAP 128

Db 60 rdgidv--yeiprlrpaqvdlpql-ltpaytyagsfvqplrtartgrprvltgldch 116

QY 129 NSVHLGSSFRRAFDGMADITEPGSFKDYVYYPNROSARTLWHDHAMHITAENARCOA 188

Db 117 anvhlngvhpatsdghpmdllppgskkydydpnlgrgatlwyhahthyeadhvryglh 176

QY 189 GLVMTDPADALNLPDSYGEFDPIMILTSKOYTANGNLVTNGELNSWGD--VIHVG 246

Db 177 gfylddpaeahlrlpa--gkydvpimlrnaqfddsgalvfghpd-----drvllang 228

QY 247 QPWPFFNVBPRKRYFRFLDAVSRSGLYFADTDAIDTFLPFVIVASDGLIEHPADTSL 306

Db 229 kagpyfevaprtyrflhlaalkhvirlnlge-----plrlatcogallpapshte 281

QY 307 LYISMAERYEVVDFSDYAGKTIELRLNGSIGIGTDYDNTDKVMRFVVAADTTOPD 366

Db 282 lalsgerveivdfaeag-----ggpv-----ylldgdnpdlldfdvsstrav-fd 326

QY 367 TSVVPANLRDVPFSPPTNTPR---QFRPGRTGPTWTINGVAF---ADVONRLLANVP 418

Db 327 psrvpvtlralp-pmgltprvterlvsfmsarplalndgkfpdlrldvq-----vk 379

QY 419 VGTVERMELINAGNG---WTHPIHILVDPFKVISRSGNNARVMEYEGSKVYVNLGR 475

Db 380 rgstelmvvnadtdpfdpnhlhlvtrlylvgigppa---pedaglkdvyspka 435

QY 476 ETVVVEAHYA-PPQVYMFCHNLIEDHDMMA 507

Db 436 gsvkigtatpylggyvynchylenslgmma 468

RESULT 15

AAAB79105
ID AAAB79105 standard. Protein: 497 AA.

XX AAAB79105;

XX 30-APR-2001 (first entry)

XX Corynebacterium glutamicum HA protein sequence SEQ ID NO:166.

KW Corynebacterium glutamicum; homeostasis; adaptation; HA protein;
 KW fine chemical production; organic acid; proteinogenic amino acid;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carboxylate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;
 KW diagnosis; Corynebacterium diptheriae; genetic engineering;
 KW Brevibacterium; environmental condition.
 XX Corynebacterium glutamicum.
 OS
 PN WO200100842-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-IB00911.
 XX
 PR 25-JUN-1999; 99US-0141031.
 PR 08-JUL-1999; 99DE-1031636.
 PR 09-JUL-1999; 99DE-1032125.
 PR 09-JUL-1999; 99DE-1032126.
 PR 09-JUL-1999; 99DE-1032127.
 PR 09-JUL-1999; 99DE-1032128.
 PR 09-JUL-1999; 99DE-1032129.
 PR 09-JUL-1999; 99DE-1032226.
 PR 14-JUL-1999; 99DE-1032920.
 PR 14-JUL-1999; 99DE-1032922.
 PR 14-JUL-1999; 99DE-1032924.
 PR 14-JUL-1999; 99DE-1032926.
 PR 14-JUL-1999; 99DE-1032930.
 PR 14-JUL-1999; 99DE-1032933.
 PR 14-JUL-1999; 99DE-1032935.
 PR 14-JUL-1999; 99DE-1032973.
 PR 14-JUL-1999; 99DE-1033002.
 PR 14-JUL-1999; 99DE-1033003.
 PR 14-JUL-1999; 99DE-1033005.
 PR 14-JUL-1999; 99DE-1033006.
 PR 31-AUG-1999; 99DE-1041378.
 PR 31-AUG-1999; 99DE-1041379.
 PR 31-AUG-1999; 99DE-1041390.
 PR 03-SEP-1999; 99DE-1041391.
 PR 03-SEP-1999; 99DE-1042088.
 XX
 PA (BADI) BASF AG.
 XX
 PI Pompejus M, Kroege B, Schroeder H, Zelder O, Haberhauer G;
 XX
 DR WPI, 2001-061974/07.
 XX
 DR N-PSDB; AAF71220.
 XX
 PT New isolated Corynebacterium glutamicum nucleic acid for production or
 PT modulation of production of fine chemicals such as amino acids,
 PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins,
 PT or enzymes -
 XX
 XX Claim 20; Page 373-374; 712pp; English.
 PS
 PS AAF71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis
 CC and adaptation (HA) proteins given in AAB79023 to AAB79242. The
 CC C. glutamicum HA genes (I) can be used in vectors for expression in host
 CC cells and production of fine chemicals, such as, an organic acid,
 CC proteinogenic or nonproteinogenic amino acid (preferred), purine or
 CC pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
 CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
 CC polypeptide or enzyme. The amino acids produced can be lysine, glutamine,
 CC glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
 CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,
 CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can
 CC be modulated. The presence of (I) or HA proteins encoded by them are
 CC used for diagnosing the presence or activity of Corynebacterium
 CC diptheriae. (I) can be used to map the C. glutamicum genome or can be
 CC used as markers for genetically engineered Corynebacterium or
 CC Brevibacterium. The HA proteins encoded by the (I) are used to maintain
 CC homeostasis in C. glutamicum or help the microorganism to adapt to

CC different environmental conditions.
 XX
 SO Sequence 497 AA:
 Query Match 16.1%; Score 494.5; DB 22; Length 497;
 Best Local Similarity 30.9%; Pred. No. 3, 2e-36;
 Matches 151; Conservative 67; Mismatches 219; Indels 51; Gaps 17;
 QY 53 LPIPPYKOPRLTYNTNVNGEITVEYEIKFTHOYPPIDGSDVLYGDSMGSPFPQYP 112
 DB 30 Irippead-----lgtregssv-hfaleaqtgseqlldv-tlktwgingthlgrlrvk 81
 QY 113 RGEYIVREFINNAEAFNSVYLHGFSRAAFDGAEDITEGSPKDYYPNROSARTLWYH 172
 DB 82 kgddvhdvdiinnidemtvtvwhgmklpaladgphbpipqgtwspwtvandaallwylh 141
 QY 173 DRAMHTTAENAVRGAGLYMLTDPADALNLPSCYGEFDIPMLITSKOYANGNLVTTNG 232
 DB 142 phthgltghayrghagmllvedeatckldlpreygvddipvlmdhrfledgsldedl 201
 QY 233 ELNSFWGDVHVHVGQFPPFNKVRPRKRRFELDAVSRSGLYFADTDAIDTRLPRKYIA 292
 DB 202 pdlgyllgdrptangitnahfdatttrvtrfvlnsgsumrftynlafsdtrt-----fgvla 255
 QY 293 SDSGLEHPADTSLAYISMAREYEVFESDYAGKTIELKNLGSGIG-GIGTD---TDYD 348
 DB 256 sdsglldepqdrtilalpggerweivlele--pgevtlesvgyfednhyvppdeftypdlg 313
 QY 349 NTDKVRFFV--ADDTTOPDTSVPANLRDVPSPPTNTPPROFRGRTGPTWTNGVA 405
 DB 314 msdsfgllltltpsdadaq--apalpgvllkstepvidatertfflm-----tfsindlq 367
 QY 406 FAVYQRLANVPVGVVERKELINAGNGWHPHILHVDKVISRSNGNARTVMFYEGS 465
 DB 368 m-dmq-rvdvdlidhdpewiivndnsdwpnfhvndarfvl-klegld---vellfndg 421
 QY 466 LKDVWL--GRETVEVE-AHYAPFGVYWFHCHNLIHEDHMAAFNMTVLDPGYCNAT 522
 DB 422 wkdtvllppataclavefghypdpqpywmhcmlyhneogumngqf----- 468
 QY 523 VFVDPME 530
 DB 469 vlivepgde 476

Search completed: October 3, 2002, 10:56:41
 Job time: 145 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 11:02:16 ; Search time 19.26 Seconds

(without alignments)
1149.926 Million cell updates/sec

Title: US-09-656-640A-4

Sequence: 1 MFKHTLGAALSLLENSNAV.....AVTERIQTMAYRPPYAADE 572

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	3067	100.0	572	1	BLRO_MYRVE
2	713	23.2	513	1	COTR_BACSU
3	523	17.1	642	1	PHSA_STRAT
4	401	13.1	516	1	YACK_ECOLI
5	315.5	10.3	470	1	SUFI_ECOLI
6	311	10.1	470	1	SUFI_SALTY
7	228.5	7.5	520	1	LACI_TRAVI
8	228	7.4	311	1	SUFI_HAEIN
9	223.5	7.3	591	1	LACI_CRYPA
10	215	7.0	520	1	LAC4_TRAVE
11	212	6.9	520	1	LAC4_TRAVI
12	210	6.8	527	1	LAC5_TRAVE
13	207	6.7	527	1	LAC5_TRAVI
14	205.5	6.7	519	1	LAC1_CORHI
15	205	6.7	519	1	LAC2_TRAVE
16	205	6.7	609	1	COPA_PESM
17	205	6.7	624	1	PER3_CANAL
18	204	6.7	519	1	LAC2_TRAVI
19	199.5	6.5	548	1	LAC1_PHLRA
20	198.5	6.5	531	1	LAC4_THACU
21	194	6.3	639	1	LAC1_NEUCR
22	193.5	6.3	622	1	FET5_YEAST
23	191	6.2	605	1	PCOA_ECOLI
24	189.5	6.2	636	1	FET3_YEAST
25	189	6.1	533	1	LAC2_PLEOS
26	187.5	6.1	621	1	LAC2_PODAN
27	185.5	6.0	529	1	LAC1_PLEOS
28	178.5	5.8	619	1	LAC2_NEUCR
29	173.5	5.7	579	1	ASO_CUCMA
30	163.5	5.3	520	1	LAC1_AGABI
31	163.5	5.3	520	1	LAC2_AGABI
32	162.5	5.3	622	1	YAR8_SCHPO
33	159.5	5.2	552	1	ASO_CUCPM

34	159.5	5.2	572	1	LAC3_THACU	002079 thanatephor
35	159	5.2	486	1	LAC1_BOMCI	012570 botrytis ci
36	150	4.9	587	1	ASO_CUCSA	P14133 cucumis sat
37	149	4.9	609	1	LAC1_EMBNI	P17489 emericella
38	143	4.7	576	1	LAC1_THACU	P56193 thanatephor
39	141	4.6	473	1	LAC3_TRAVI	099049 trameetes vi
40	134	4.4	599	1	LAC2_THACU	020075 thanatephor
41	131	4.3	462	1	CBSA_SULAC	054088 sulfolobus
42	113.5	3.7	578	1	ASO_TORAC	040588 nicotiana t
43	108.5	3.5	781	1	GCST_CADEL	019426 caenorhabdl
44	108	3.5	746	1	ABP_RAT	P36633 rattus norv
45	108	3.5	865	1	LOX2_ORYSA	P29250 oryza sativ

ALIGNMENTS

RESULT	ID	BLRO_MYRVE	STANDARD;	PRT;	572 AA.
AC	012737,				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	Bilirubin oxidase precursor (EC 1.3.3.5).				
OS	Myrothecium verrucaria.				
OC	Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Myrothecium.				
OX	NCBI_TaxID=5532;				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	SPRAIN-MT-1;				
RX	MEDLINE=9336794; PubMed=8360171;				
RA	Koike S., Ando K., Kaji H., Inoue T., Murao S., Takeuchi K., Samejima T.;				
RT	Molecular cloning of the gene for bilirubin oxidase from Myrothecium verrucaria and its expression in yeast.;				
RL	J. Biol. Chem. 268:18801-18809(1993).				
CC	- FUNCTION: OXIDATION OF BILIRUBIN AND OTHER TETRAPYRROLES.				
CC	- CATALYTIC ACTIVITY: Bilirubin + O(2) = biliverdin + H(2)O.				
CC	- COPFATOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH CONTAINE THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR. CONTAINS 2 BLUE COPPER ATOMS PER MOLECULE.				
CC	- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.				
CC	- SIMILARITY: CONTAINS 2 PLASTOCYANIN-LIKE DOMAINS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: D14081; BAA03166.1; -				
DR	EMBL: D12579; BAA02123.1; -				
DR	InterPro: IPR001117; Cu-oxidase.				
DR	Pfam: PF00394; Cu-oxidase; 1.				
KW	Signal; Copper; Metal-binding; Oxidoreductase; Glycoprotein; Repeat.				
FT	SGNML	1	19		PROBABLE.
FT	PROPEP	20	38		
FT	CHAIN	39	572		BILIRUBIN OXIDASE.
FT	DOMAIN	98	194		PLASTOCYANIN-LIKE 1.
FT	DOMAIN	404	526		PLASTOCYANIN-LIKE 2.
FT	METAL	132	132		COPPER (TYPE 2) (BY SIMILARITY).
FT	METAL	134	134		COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	172	172		COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	174	174		COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	436	436		COPPER (TYPE 1) (BY SIMILARITY).
FT	METAL	439	439		COPPER (TYPE 2) (BY SIMILARITY).
FT	METAL	441	441		COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	494	494		COPPER (TYPE 3) (BY SIMILARITY).
FT	METAL	495	495		COPPER (TYPE 1) (BY SIMILARITY).

FT METAL 496 496 COPPER (TYPE 3) (BY SIMILARITY).
 FT METAL 500 500 COPPER (TYPE 1) (BY SIMILARITY).
 FT METAL 503 503 COPPER (TYPE 1) (BY SIMILARITY).
 FT CARBOHYD 510 510 N-LINKED (GLCANC. . .) (POTENTIAL).
 FT CARBOHYD 520 520 N-LINKED (GLCANC. . .) (POTENTIAL).
 SO SEQUENCE 572 AA: 63947 MW: 58426413035E9FF CRC64:

Query Match 100.0%; Score 3067; DB 1; Length 572;
 Best Local Similarity 100.0%; Pred. No. 4.3e-214;
 Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHTLGAALSLLENSNAVQASVPETSPATGHLKRYAQISPOYPTFTVLPPIPVQ 60
 DB 1 MKHTLGAALSLLENSNAVQASVPETSPATGHLKRYAQISPOYPTFTVLPPIPVQ 60
 QY 61 PRLTYNPNNGEIMWEYVEIKPFTHQVYPDLSADLVGDSNPEPTQVPRGVETVVR 120
 DB 61 PRLTYNPNNGEIMWEYVEIKPFTHQVYPDLSADLVGDSNPEPTQVPRGVETVVR 120
 QY 121 FINNAEAPNSVHLHGSFSAFADGMAEDITEPGSEKDYVYYPNRSARTLMYHDHAMHTA 180
 DB 121 FINNAEAPNSVHLHGSFSAFADGMAEDITEPGSEKDYVYYPNRSARTLMYHDHAMHTA 180
 QY 181 ENAYRQAGLYMLTPDAEDALNPSGYGFDIPMLITSQYRANGMLVTTNGELNPFMD 240
 DB 181 ENAYRQAGLYMLTPDAEDALNPSGYGFDIPMLITSQYRANGMLVTTNGELNPFMD 240
 QY 241 VHVNGQPPFPKVRFRKFRFLDAASRSGLYPADDAIDTRLPFYKIASDSGLLEH 300
 DB 241 VHVNGQPPFPKVRFRKFRFLDAASRSGLYPADDAIDTRLPFYKIASDSGLLEH 300
 QY 301 PADTSLLYISMAERYEVDFSDYAGKTLELRNLGSGIGIGTDYDNTDKVMRFVAD 360
 DB 301 PADTSLLYISMAERYEVDFSDYAGKTLELRNLGSGIGIGTDYDNTDKVMRFVAD 360
 QY 361 DTTQPTSVVPAALRVPPSPPTTNPFRGRTGPTMTINGVAAVONRLAVPVG 420
 DB 361 DTTQPTSVVPAALRVPPSPPTTNPFRGRTGPTMTINGVAAVONRLAVPVG 420
 QY 421 TVERMELIANGQWPHIHLVDFKVISRTSGNNARTVAPYSGLDKDVYLGRRRETVV 480
 DB 421 TVERMELIANGQWPHIHLVDFKVISRTSGNNARTVAPYSGLDKDVYLGRRRETVV 480
 QY 481 EAHYAFEPGVYMFHCHNLJHEDHMAAFNATVLPYGINATVFPDMEELQARPEYLG 540
 DB 481 EAHYAFEPGVYMFHCHNLJHEDHMAAFNATVLPYGINATVFPDMEELQARPEYLG 540
 QY 541 EFOAGSGFSVQAVTERIOTMAEYRYAADE 572
 DB 541 EFOAGSGFSVQAVTERIOTMAEYRYAADE 572

RESULT 2
 COFA_BACSU STANDARD: PRT: 513 AA.

ID COFA_BACSU 024818;
 AC P07788; 024818;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Spore coat protein A.
 GN COFA OR PIG.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124186; PubMed=8969499;
 RA Boriss R., Portollik S., Schreier R.;
 RT "The 52 degrees-55 degrees segment of the Bacillus subtilis
 chromosome: a region devoted to purine uptake and metabolism, and

RT containing the genes cotA, gapB and guaA and the pur gene cluster
 RT within a 34960 bp nucleotide sequence.";
 RL Microbiology 142:3027-3031(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / MARBURG;
 RX MEDLINE=98116660; PubMed=9455482;
 RA Kasahara Y., Nakai S., Ogasawara N., Yata K., Sadate Y.;
 RT "Sequence analysis of the groESL-cotA region of the Bacillus subtilis
 genome, containing the restriction/modification system genes.";
 RL DNA Res. 4:335-339(1997).
 RN [3]
 RP SEQUENCE OF 1-37 FROM N.A.
 RX MEDLINE=88011308; PubMed=2821284;
 RA Donovan W., Zheng L., Sandman K., Losick R.;
 RT "Genes encoding spore coat polypeptides from Bacillus subtilis.";
 RL J. Mol. Biol. 196:1-10(1987).
 RN [4]
 RP SEQUENCE OF 1-34 FROM N.A.
 RX MEDLINE=88286730; PubMed=3135411;
 RA Sandman K., Kroos L., Cutting S.M., Youngman P., Losick R.;
 RT "Identification of the promoter for a spore coat protein gene in
 RT Bacillus subtilis and studies on the regulation of its induction at a
 RT late stage of sporulation.";
 RL J. Mol. Biol. 200:461-473(1988).
 RN [5]
 RP SEQUENCE OF 1-10 FROM N.A.
 RC STRAIN=168;
 RA Wray L.V., Ferson A.E., Fisher S.H.;
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN BROWN PIGMENTATION DURING SPOROGENESIS.
 CC -1- SIMILARITY: TO S.ANTIBIOTICUS PHENOXAZINONE SYNTHASE (PHSA).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U51115; AAB62305.1;
 DR EMBL: AB007636; BAA22774.1; ALT_INT.
 DR EMBL: 299107; CAB12449.1;
 DR EMBL: X05678; CAA29165.1; ALT_INT.
 DR EMBL: X07512; CAA50392.1;
 DR EMBL: U31756; AAC44642.1;
 DR PIR: A27393; A27393.
 DR Subtilist: BG10490; cotA.
 DR InterPro: IPR001117; Cu-oxidase.
 DR Pfam: PF00394; Cu-oxidase; 1.
 KW Sporulation; Complete proteome.
 FT CONFLICT 347 367 DESRRKPLASYPVSQVHERIQ -> TKAESRSTSPHTLTKS
 FT CONFLICT 451 458 GPVAVPPPP -> VRCFAAA (IN REF. 1).
 FT CONFLICT 451 458 GPVAVPPPP -> VRCFAAA (IN REF. 1).
 SO SEQUENCE 513 AA: 58499 MW: 8368345807587 CRC64;

Query Match 23.2%; Score 713; DB 1; Length 513;
 Best Local Similarity 33.1%; Pred. No. 3.5e-44;
 Matches 177; Conservative 77; Mismatches 179; Indels 102; Gaps 16;

QY 49 FTVPLPDP---PVKQRLTVTPVNGQELWYVEIKPFTHQVYPDLSADLVGDSN 104
 DB 6 FVDALPDPDLKPVQO-----SKETYYEVMECTHQLHRDLPTRLMGVNGLF 55
 QY 105 GGRTEQVGRGVETVVRINN-----AEAPNSVHLHGSFSAF 142
 DB 56 GGRTEQVGRGVETVVRINN-----AEAPNSVHLHGSFSAF 142
 QY 143 DGMAE-----DTEPGSF--KDYVYYPNRSARTLMYHDHAMHTENAYRQAGLYMLT 194
 DB 143 DGMAE-----DTEPGSF--KDYVYYPNRSARTLMYHDHAMHTENAYRQAGLYMLT 194

DB 116 DGYPEAMFESKDEPQGTGTYGKREYVYHPNQGALIMYHDMALITLNVYAGLVAYIHH 175
QY 195 DPAEDALNIPSGYGEFDIPMLITSKOYTANGNLVTNGELN-----SFMGDVI 242
DB 176 DPEKRLKLPKPS--DEYDYLPLITDRTINEGSLFYPSPAPNPSPSLPNPSTVAFCCETI 233
QY 243 HVGQGPWPKFKNVPRKRYREFELDAVSRSGFLYFADTDALITLPRKVIASDGLLEHFA 302
DB 234 LVNGKWMPYLEVPEPRKRYFRVIVNASMTRYNL-----SLDNGGDFLQIGSDGGLLPRSV 287
QY 303 DTSLLITISMAEREVYVDFSDYAGKITELRNLGSGISGIGITDIDYDNTKVMKFAVDAT 362
DB 288 KINSLFSLAPAEKRDIIIDIFAYEGESIIILANSAGCGDVNPEDAN---IMQFRVTKPL 343
QY 363 TOPDTSVPAKLDVPPSPPTTNTPROF-----FGRTGPTWTINGVAFADVQNR 412
DB 344 AOWDESKRKRYL--ASYPSVQHERIQNIRKLKLAGTODEGR--PULLLNKKRWHDP--- 336
QY 413 LIANVP-VGTEVERMELINAGNGWTHPIHLVDPKYISRTSGNNAR----- 457
DB 397 -VTEPKVGTETIWSIINPTRG--THPIHLVSPFVLDRRPFIDARVQESGELSYTGPAV 454
QY 458 TWMPYFESGLKDVWVIGRRETVVEAHYAPPYGMCHNLTHEDHMAAFNAT 512
DB 455 PPPSEKGMKDTIQAHAGEVLRITAAETFGYSGRYVWHCHLHEDYDMRPMNDIT 509
RESULT 3
PHSA_STRAT STANDARD: PRT: 642 AA.
ID PHSA_STRAT 053692;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Phenoxazinone synthase (EC 1...5) (PHS).
GN PHSA.
OS Streptomyces antibioticus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OY NCBI_TaxID=1890;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RC STRAIN=IMRU 3720;
RX MEDLINE=9601355; PubMed=7592317;
RA Hsieh C.-J., Jones G. H.;
RT "Nucleotide sequence, transcriptional analysis, and glucose
regulation of the phenoxazinone synthase gene (phs) from
Streptomyces antibioticus."
RT J. Bacteriol. 177:5740-5747(1995).
RN [2]
RP SUBUNIT.
RX MEDLINE=82066838; PubMed=7305384;
RA Choy H.A., Jones G.H.;
RT "Phenoxazinone synthase from Streptomyces antibioticus: purification
of the large and small enzyme forms."
RT Arch. Biochem. Biophys. 211:55-65(1981).
CC -1- FUNCTION: CATALYZES THE LAST BUT TWO STEPS IN THE PUTATIVE
BIOSYNTHETIC PATHWAY OF ACTINOMYCIN.
CC -1- CATALYTIC ACTIVITY: 4 4-methyl-3-hydroxyanthraniloyl pentapeptide
+ 3 O(2) = 2 actinomycinic acid + 6 H(2)O.
CC -1- COFACTOR: THIS PROTEIN BELONGS TO THE MULTICOPPER OXIDASES WHICH
CONTAIN THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE
2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (SMALL FORM) OR HOMOHXAMER (LARGE FORM).
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 2 PLASTOCYANIN-LIKE DOMAINS.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
CC EMBL: U04283; AA86668.1; ALT_INIT.
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; Multicopper oxidase2.
DR Pfam: PF00394; Cu-oxidase; 1.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
KW oxidoreductase; Repeat; Metal-binding; Copper;
KW Antibiotic biosynthesis.
FT INIT_MET 0
FT DOMAIN 86 223 PLASTOCYANIN-LIKE 1.
FT METAL 493 620 PLASTOCYANIN-LIKE 2.
FT METAL 161 161 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 163 163 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 201 201 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 203 203 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 524 524 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 527 527 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 529 529 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 602 602 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 603 603 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 604 604 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 608 608 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 613 613 COPPER (TYPE 1) (BY SIMILARITY).
SQ SEQUENCE 642 AA; 70113 MW; E0B39C0BA3364E48 CRC64;

Query Match 17.18; Score 523; DB 1; Length 642;
Best local similarity 27.8%; Pred. No. 2.5e-30;
Matches 176; Conservative 83; Mismatches 199; Indels 174; Gaps 28;

QY 27 ETSPATGHLFKRAQISPOQPMFTVPLPIPPVQOPRLTVNPNVNGOITWYEVEIKRFTH 86
DB 33 EQAPAPQ-----ELTP-----FAAPLTVPVLP--ASDEVTRT-----ETALRPTWV 74
QY 87 QVVPDLGSADLVGYDGKSPGPTFOVPGVETVVRFIN-----NAE 126
DB 75 RLHPQLPPTLMGVDGVPPTTEVRGQVRIAMWIRIKGSEIYPTSVELGPPGTP 134
QY 127 APNS-----VHLGSFSAADFGMAEDITTEPSFKDYYPNROS 165
DB 135 APMTPEKGGVEPNKDVAALPAMSVTHLGAQTGGNDGMADNAVGGDAQLEYPNDHQ 194
QY 166 ARTLWYHDMAMHTTAENAYGQAGLWLTDPADALNIPSGYGEFDIPML-----TSK 219
DB 195 ATQWYHDMAMITRNMVMAGLVGYLVLRDDEBDALGLPS--GDREIPLLIADRNLDITDE 252
QY 220 QYRANGNLY-----TTNGELMS-----FMGDVYHNGSPWPKFKNVPRKRYREFLDA 267
DB 253 DGRNLGRHLKRYIVQGSNPEETKRPVSIPEFGYTYTVNGRIMPYADDDGWYRLRVNAS 312
QY 268 VSRSGFLYFADTDALITLRFKY--IASDSGLEHFA-----DT-SLITSMAEREVEVF 319
DB 313 NARTVNLVLDED--DRPVGVVHQIGSDGGLLPRVPVDFDTLPLSAAAPRERDLIV 370
QY 320 DESDYAGKITELNLC-----GSIGIGITPTDIDNTDKVMKFAVDATDTPDTSVVP 371
DB 371 DFRALGGRRLRLVDKGGAPAGTPDPLGAV-----RYPEVMEFRVR-ETCEDSDPALP 422
QY 372 ANL-----RDVPF-----PSPTNT-----PROPRF----- 392
DB 423 EVLSGSRMSHDIPIGHRILIVLTPEGTGSGGHPETIWEAEVEDPADVOVPAEGVIQVT 482
QY 393 GRTGPTWTINGVAFADVQNRLLANVPVGTVERMELINAGNGWTHPIHLVDPKYISR-- 450
DB 483 GAGGRKTYRRTA-ATPNDSLGFTIGSGTHQWTFNLN-SPILHPHIIHLADQVIGRDA 540
QY 451 --TSG-----NART-----VMYFESGLKDVWVIGRRETVVEAHYAPPYGM 492
DB 541 YDASGFDALGTRTVRLDPDTPVPLAPNELCHKIDVFQVPGQGLRVMGKFDGAGVGRFM 600
QY 493 FHCNHLIHEDHMAAFNATVLP-----DYG 518

OK NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Berna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirovpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE OF 1-89 FROM N.A.
 RX MEDLINE=92212294; PubMed=1557036;
 RA Coleman J.;
 RT "Characterization of the Escherichia coli gene for
 RT 1-acyl-sn-glycerol-3-phosphate acyltransferase (plac).";
 RL Mol. Genet. 232:295-303(1992).
 RN [3]
 RP SEQUENCE OF 28-39.
 RC STRAIN-K12 / EMG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 RT in the genome of Escherichia coli K-12.";
 RL Electrophoresis 18:1259-1313(1997).
 CC -1- FUNCTION: INVOLVED IN CELL DIVISION, SUPPRESSES A FTSI MUTATION.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U28377; AAA69185.1; -
 CC EMBL: AE000384; AAC76053.1; -
 CC EMBL: M63491; AAC24398.1; -
 CC PIR: S20461; S20461.
 CC Ecogene: EG11376; sufi.
 CC InterPro: IPR001117; Cu-oxidase.
 CC Pfam: PF00394; Cu-oxidase; 1.
 CC DR Periplasmic; Signal: Complete proteome.
 CC FT SIGNAL 1 27
 CC FT CHAIN 28 470 PROTEIN SUFI.
 CC FT DOMAIN 68 164 PLASTOCYANIN-LIKE.
 CC FT SEQUENCE 470 AA; 51858 MW; C843A5A4CB146688 CRC64;
 SQ
 Query Match 10.3%; Score 315.5; DB 1; Length 470;
 Best Local Similarity 25.9%; Pred. No. 1.6e-15;
 Matches 133; Conservative 71; Mismatches 235; Indels 75; Gaps 21;
 Oy 19 AVOASVPEPTSPATGHLFRVAQISQVPMFYVPIIPVVKORLVNTPVNGOELIWEY 78
 Db 15 ALCAGAVPLKASAGQ-----QQPLVPPLLEER-----RGQPL-FMT 51
 Oy 79 VEIKPTHOVYDPLGSAADLVGYDGMSPGPTFOVPRGCVETVREINNAEPNSVHLGSFS 138
 Db 52 VQ---RAHMSFPGTRASVWINGRYLGPTRIRWKGDDVKLIYSNRLTENVSMTVAGLQY 108
 Oy 139 RAAFGWADIEPQSFKRYIYPNROSARTLWYHDHAMITAEANVRGQAGLYMLTDPAR 198
 Db 109 PGRLMGAPRPMSPNDMAVPLPIRQNAATLWYHANTPRTAQOYVNGLAGMVLDEVS 168
 Oy 199 DALNPSGGEEDIPMILSKOTYTAGNLYTTNGELNSFMGDVHYNGQDPWPKVNEPRK 256
 Db 169 KSLPIFNHGVDDFYIIODKRLDNGCTPEYNEPSSGSGVGGDTLLVNGQSPVVEYSRGM 228
 Oy 259 YRFRFLDAVSRSGFLYFADTDAIDTRLPFKVYIASDSGLLEHDPDTSLLYISMAERYEVV 318

Db 229 VRLRLNLSNRRYQLOQND-----GRPLHVISGQGFPAVSVKQSLACGERREIL 282
 Oy 319 FDFSVDYAGKTELRNLGSGTIGCTDTDYNTDKVAFVYADD-----TQOPTSVY 370
 Db 283 VDSN--GDEVSI--TCGEASTI-----VDRIKGFEPSSILVSTVLTLRP-TGL 329
 Oy 371 PAMLRVPPF-SPT---TNPFRGFRGTCGPTWTINGVAPADVQNRLLANVPYGVVERME 426
 Db 330 PLVTDSLPMKPLFTELMAGSPITSRDISLGDDEINGQLM-DV-NRIDVTAQOGTWERW- 386
 Oy 427 LINAGMGWTHPIHILVDFEKVVISRTSGNNARIVMPY--ESGLDYYWL-GRETVVV--E 481
 Db 387 TVRADE--PQAFHIEGVMEQI-----RNVNGAMPFEDRGMDVWVGQVELLYVFGQ 438
 Oy 482 AHYAPFGYVFMFCHNLHEDHDMMAAFNATVP 515
 Db 439 PSMHFP--FYFNSQTLFEMADRGISQGLLVNPPV 470
 RESULT 6
 SUFI_SALTY STANDARD: PRT; 470 AA.
 ID SUFI_SALTY
 AC P40799;
 DT 01-FEB-1995 (rel. 31, Created)
 DT 01-MAR-2002 (rel. 41, Last sequence update)
 DT 01-MAR-2002 (rel. 41, Last annotation update)
 DE Protein sufi precursor.
 GN SUFI OR STM3172 OR STY3349.
 OS Salmonella typhimurium, and
 OS Salmonella typhimurium, gamma subdivision; Enterobacteriaceae;
 OC Bacteria; Proteobacteria;
 CC Salmonella.
 CC NCBI_TaxID=602, 601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. typhimurium; STRAIN=LT2;
 RA Cong J., Schmid M.B.;
 RL Submitted (APR-1994) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. typhimurium; STRAIN=LT2 / SGC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. typhimurium; STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jags K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar typhimurium CT18.";
 RL Nature 413:848-852(2001).
 CC -1- FUNCTION: INVOLVED IN CELL DIVISION, SUPPRESSES A FTSI MUTATION
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 413.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: U09309; AAL56679.1; ALT_FRAME.
DR EMBL: AEO08845; AAL22046.1; -
DR EMBL: AL627277; CAD03004.1; -
DR StyGene: SG10522; sufl.
DR InterPro: IPR001117; Cu-oxidase.
DR Pfam: PF00394; Cu-oxidase; 1.
KW Periplasmic; Signal; Complete proteome.
FT SIGNAL 1 BY SIMILARITY.
FT CHAIN 28 470 PROTEIN SUFL.
FT DOMAIN 68 164 PLASTOCYANIN-LIKE.
SQ SEQUENCE 470 AA; 51858 MW; 61E84D4D42B025FB CRC64;

Query Match 10.1%; Score 311; DB 1; Length 470;
Best Local Similarity 25.4%; Pred. No. 3,4e-15;
Matches 121; Conservative 65; Mismatches 206; Indels 84; Gaps 18;

OY 52 PLIPVVKOPRLTVNPNVNGEIMYEVEIKPFTHQVYVYDLGSAADLVGSGMSPGTFQV 111
DB 33 PLVPRPLESR-----RCQPLFM---TLQRAHMSFTQGRAPAVGCVNGVYLGPTIRV 31
OY 112 PRCVETVVRFINNAEPNSVHLHGSFSRAAFDGMADITEPGSEKDYVYVYVYVYVYV 171
DB 82 WKGDVVKLLYSNRLAENVMSTVAGLLVPGRLMGCPARMSPNADAPVPIRISOATLWY 141
OY 112 HDHAMHTAENMVRQAGLYMLTDAEDALNPSGGEEDIMILTSKQYANGNLYVTN 231
DB 142 HANPNKRAQOYVYNGAGWMLVEDDISKTLPIPNHGVDDFVITQDKLDFNGPVESE 201
OY 232 GELNSFWGQVIVHNGOPMFEKVRERPRKYRFRFLDAVSRSGFLYVADTDALTRLPFKYI 291
DB 202 PGSGFVGDTLLVNGAQSRYEVESRGWVRLRLNMSNRQYLOMGSDGAL-----HVI 255
OY 292 ASDSGLEHPRADTSLIYSMAERIEVDFSDYACKTIRLNLGSGIGTIDTDYDNTD 351
DB 256 SGDOGFLEPAPVSVKOLSLAPGERREILVDMTN--GDEVSI--TCEPAASI-----VD 303
OY 352 KVMREYVADDDTQDPDSVY-----PANLRDVPPEPSTPTNPRQFRGRTGPTMTING 403
DB 304 RIRGTF-----EPSSILVSTLVTLRPTGL---LPLVTDNLPW-----RLPPEIMSG 348
OY 404 --VAFAVQ-----NRLANVPVGTVERMELINAGNGWTHPIHILVDFKV 447
DB 349 APVRSRDISLGDDPGINGLMDVNRIDITAOQGTWERM---TVRADMPGSPHIEGVSFLI 405
OY 448 ISRTSGNNAKRYMPY--ESGLKDVVWL--GRRETVVY--EAHAPPGVYMFCHNL 498
DB 406 -----RVNNGAMPPEPDRGMDTVWDQVELLVYVYGOPSPWHPF--EYFNSQTL 453

RESULT 7
LAC1_TRAVI STANDARD; PRT; 520 AA.
AC Q99044;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Laccase I precursor (EC 1.10.3.2) (Benzenediol: oxygen oxidoreductase)
DE (Urishiol oxidase) (Diphenol oxidase).
GN LCC1.
OS Trameetes villosa (White-rot fungus).
OS Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticiaceae; Trameetes.
OX NCBI_TaxID=47662;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Mycelium;
RC MEDLINE=97076915; PubMed=8975613;
RA Xaver D.S., Xu F., Colightly E.J., Brown K.M., Brown S.H.,
RA Ray M.W., Schneider P., Halkier J., Mondorf K., Dalboge H.,
RT "Purification, characterization, molecular cloning, and expression of
RT two laccase genes from the white rot basidiomycete Trameetes
RT villosa".
RL Appl. Environ. Microbiol. 62:834-841(1996).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DEOXYIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLE).
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzoquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: L49376; AAC1686.1; -
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; MultiCu-oxidase2.
DR Pfam: PF00394; Cu-oxidase; 3.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1.
DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; FALSE_NEG.
KW Oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;
KW Lignin degradation; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 520
FT DOMAIN 23 148
FT DOMAIN 160 302
FT DOMAIN 369 491
FT METAL 85 85
FT METAL 87 87
FT METAL 130 130
FT METAL 132 132
FT METAL 416 416
FT METAL 419 419
FT METAL 421 421
FT METAL 473 473
FT METAL 474 474
FT METAL 475 475
FT METAL 479 479
FT METAL 7 7
FT CARBOHYD 72 72
FT CARBOHYD 75 75
FT CARBOHYD 229 229
FT CARBOHYD 238 238
FT CARBOHYD 354 354
FT CARBOHYD 361 361
FT CARBOHYD 457 457
SQ SEQUENCE 520 AA; 55545 MW; 046AB6D74737C60E CRC64;

Query Match 7.5%; Score 228.5; DB 1; Length 520;
Best Local Similarity 22.8%; Pred. No. 3.5e-09;
Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

OY 55 IPVPVKOPRLTVNPNVNGEIMYEVEIKPFTHQVYVYDLGSAADLVGSGMSPGTFQVPRG 114
DB 23 IGPVAD--LITRNA-----VSPDGSROAVYVNGTPEPLITGNMG 62
OY 115 VETVVRFINNAE-----APNSVHLHGSFSRAAFDGMADITEPGSF-KDYVY 160

```

Db      63 DRPOLYVIDMLNHTWVKSSTSHMHGFFOKGT--NMADGPAFTNOCPISSGSHFLYDFQV 120
OY      161 PNBOSARTLWHDHAMHITAEANAYRGAGLYMLTDPEDALNIPSGGEEDIMILTSKQ 220
Db      121 PDO--AGTFYHSH--LSTQYCDLGRGFVYVDPNDPADLYVDN-----DDVTTLVDW 172
OY      221 YTAGNGLVTTNGELNSFWGDVIVNGQ-----PWPKNVEP-RKYRFRFLDAVSR 270
Db      173 YHAAKL-----GPAFLGADATLILNGKRSPTTADLSVISTPKRYRFLVLSLSDP 228
OY      271 SF-----GLYPADDAIDTRLPEKVIASDSGLLEHPADTSLIYSMAEREVEFEDSD 333
Db      229 NYTFSDIGHNMFTIETDSINTA-----PLVVDSTQIPAAQRYSEVLE-AN 272
OY      324 YAKTIELR---NG--GSIGIGITDIDYNTDKVMF---VYADDTQDDTSVPANLR 375
Db      273 QAVDNTWIRANPNFNGVFTGINS-----ALTRYDGAALVEPTTQTSTAPINEV 324
OY      376 DVEFPSTPTTPR-----OEFGRGPTWTINGVAFADVONRLLANVPGTVER 424
Db      325 NL-HPLVTTAVPGSPVAGVDLAIMNAFNNGTNEFINGSTFPPYVPLQIISGQNA 383
OY      425 WEIINAGN-----GWHPRIHILVDFKVIYSTSGNNAKTYMPE 463
Db      384 QDLIPSGSVYSLPSNADIEISFPATAAPGAPHPFHILGHAFVY--RSAGS---TYVNYD 439
OY      464 SGL-KDVMYMG---RRETVYEAHYAFPGVYMFCHNLHEDHDMMAANATVLPYGY 519
Db      440 NPFRDVSSTGTAAAGDNTTIRF-TDNPGRMFLCHIDFHLGPAVVE-AEDIPD--- 494
OY      520 NATVEVDPMELW 532
Db      495 --VASANPVQAW 505

```

RESULT 8

```

SUFL_HAEIN
ID SUFL_HAEIN STANDARD; PRT; 311 AA.
AC P4847:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein sufl homolog precursor.
SUFL OR HI0733.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips R., Liu L.-I., Glodek A., Kelley J.M.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon K.C.,
RA Fine L.D., Fritchman J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC -1- SIMILARITY: STRONG, TO E.COLI AND S.TYPHIMORIUM SUFL.
CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32756; AAC22390.1; -.
KW TIGR: H10733; -.
KW periplasmic; signal; Complete proteome.
FT SIGNAL 1 28 BY SIMILARITY.
FT CHAIN 29 311 PROTEIN SUFL HOMOLOG.
FT DOMAIN 68 164 PLASTOCYANIN-LIKE.
SQ SEQUENCE 311 AA; 34496 MW; 436173F59F1E1772 CRC64;

```

Query Match 7.4%; Score 228; DB 1; Length 311;
 Best Local Similarity 25.9%; Pred. No. 1; 9e-09;
 Matches 71; Conservative 50; Mismatches 135; Indels 18; Gaps 5;

```

OY 50 TVPLPIPVKOPRLTYTPNV--NGOEIWEVEIKPFTHQVYVPLDGLSADLV--GYDGS 104
Db 21 TVPAPLLAASREKLVVPLIEVRGRPI-----VLTMOETNYPLDSSHVTVWVGFGNXY 74
OY 105 PCPFPQVPRGEVYVRRINNAEAPNSVHLGSSKRAFDGMADITPESFKDYIYPNQ 164
Db 75 LGPTIKIKSGSFAKLNHNLPPOSVALSIGLQASGEFLFGAARVLKKGSMAPVPIEQ 134
OY 165 SARTLWYHDAMHITAEANAYRGAGLYMLTDPEDALNIPSGGEEDIPMITRSKQYAN 224
Db 135 PAASCWRSATLANSAYOTRGLAGMWLIDEOSKANLPKKGCVDDITPLIDMERNND 194
OY 225 GNLVTTNGELNSFWGDVIVNGQWPPEKVEPRKRFRLDAVSRSGLYPADTAIDT 284
Db 195 GLQLEKONQPH-FVGNRLNLVNGIEAPYLDVARGMIRLRILNASLARAYDL-----RLDN 247
OY 285 RLEPKVIASDSGLLEHPADTSLIYSMAEREYV 318
Db 248 DQEMLLTAQDLGFLPKRAKSVKSLVSPGERAEIL 281

```

RESULT 9

```

LACL_CRYPA
ID LACL_CRYPA STANDARD; PRT; 591 AA.
AC Q03966;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Laccase precursor (EC 1.10.3.2) (Benzenediol: oxygen oxidoreductase)
DE (urishiol oxidase).
DE IAC-1.
OS Cryphonectria parasitica (Chesnut blight fungus) (Endothia
OS parasitica).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Diaporthales; Valsaceae; Cryphonectria.
OX NCBI_TaxID=5116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92314427; PubMed=1535523;
RA Choi G.H., Larson T.G., Nuss D.L.;
RT "Molecular analysis of the laccase gene from the chestnut blight
RT fungus and selective suppression of its expression in an isogenic
RT hypovirulent strain.";
RL Mol. Plant Microbe Interact. 5:119-128(1992).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLE).
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzenesemiquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: secreted (potential).
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


FT DISULFID 139 227 PROBABLE.
FT METAL 86 86 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 88 86 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 131 131 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 133 133 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 418 418 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 421 421 COPPER (TYPE 2) (BY SIMILARITY).
FT METAL 423 423 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 473 473 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 474 474 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 475 475 COPPER (TYPE 3) (BY SIMILARITY).
FT METAL 479 479 COPPER (TYPE 1) (BY SIMILARITY).
FT METAL 7 7 COPPER (TYPE 1) (BY SIMILARITY).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 520 AA; 55643 MW; FC1032400E9DDEA4 CRC64;

Query Match 7.0%; Score 215; DB 1; Length 520;
Best Local Similarity 21.5%; Pred. No. 3.3e-08;
Matches 119; Conservative 60; Mismatches 219; Indels 156; Gaps 23;

QY 34 HLEKRAQISPOYPMFTVLPPIPKQRLTVNPNGOEIWEYEIKPFTHQVYEDLG 93
15 HSRGRVS-----AALGPVTD-LTISNA-----DVSPDGF 42
QY 94 SADLVGYDOGSGPPTQVPRGVTYVFINNAE-----APNSVHLGSGSRAAFDGMAE- 147
43 TRAAVLAVGVPRLTLTGKGNFQINVIDNLSEMTLKSTSIHMGHGFQKGT--WMADG 100
QY 148 -----DTEPGESEKDYVYNNQASRLMYHDHMHITAEKNYRQAGLYMTDRAEDA 200
101 AAFVNOCPATGNSFL-YDPTATDQGTWYHSH-LSTQYCDGLRPMVYVYDPSDPAHL 158
QY 201 LNLPSGYGEFDIPMLITSKQYANGNL-----VTNGELNSFGWD-----VIHV 244
159 YDV-----DDEPTIILSDWYHTAASLGAAPITGSDSTILNG-LGRFAGGDSIDLAVIY 212
QY 245 NGQPMFKVNEPRKYRFRLLDAVSRSF-----GLYFADTDAIDTRLPEKVIASDSGL 297
213 E-----QGRYRMRLLISCDEPNYVFSIDGHNMTIIEADAVN----- 249
QY 298 LEHPADTSLIXYSMARVEYV-----DFSDYAKTIELNLGSGISGIGDTDYDNTDKY 353
250 -HEPLTVDSIOIYAGORYSFVLTAODDIDNYETRALPSAGTTSFSDGINS-----AI 300
QY 354 MREYVA---DDETPQDPTSVV---ANLRDVPFSPPT-----TNPQGFGRGTGPT 398
301 LKTSQSEVDPTTETTSVLPLDEANL--VPLDSPAPAGDPNIGVDYALNLDNFNDGIN 358
QY 399 WTINGVAFDVONRLANVPCTVERWELINAGN-----GWNH 436
359 FEINDVSFSPYVLLQILSGTSSAADLLPSGSLFALPSNSTIETISFPIITATNAGAPH 418
QY 437 PIHIIHVDKVISRISGNNARIVMESEGLKQYVMLGRETYVVEAHAPFCGVYVHCH 496
419 PHLHGHFHSIVRTAGSTJFNVNPR---RDVYVNTGTAGDVNTIRPTDNGPFWLHCH 475
QY 497 NLIHEDHDMMAFN 510
DB 476 IDPHLEAGFAIVFS 489

RESULT 11
LAC4_TRAVI STANDARD; PRT; 520 AA.
AC 09055:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Laccase 4 precursor (EC 1.10.3.2) (Benzenediol: oxygen oxidoreductase)
DE (Urishiol oxidase) (Diphenol oxidase).
GN ICC4.
OS Trametes villosa (White-rot fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticolaceae; Trametes.
NCBI_Taxid=47662;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97128774; PubMed=8973314;
RA Taver D.S., Gollightly E.J.;
RT Cloning and characterization of three laccase genes from the
RT white-rot basidiomycete Trametes villosa: genomic organization of the
RT laccase gene family.";
RL gene 181:95-102(1996).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLE).
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzenesemiquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: L78077; AAB47734.1; -
CC HSSP: P37064; IAOZ.
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR002355; Multicopper oxidase.
DR Pfam: PF00394; Cu-oxidase; 3.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; FALSE NEG.
KW oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;
KW Lignin degradation; Multigene family.
FT SIGNAL 1 18
FT CHAIN 19 520
FT DOMAIN 24 149
FT DOMAIN 161 303
FT DOMAIN 370 491
FT METAL 86 86
FT METAL 88 88
FT METAL 131 131
FT METAL 133 133
FT METAL 418 418
FT METAL 421 421
FT METAL 423 423
FT METAL 473 473
FT METAL 474 474
FT METAL 475 475
FT METAL 479 479
FT METAL 7 7
FT METAL 73 73
FT CARBOHYD 76 76
FT CARBOHYD 239 239
FT CARBOHYD 399 399
FT CARBOHYD 457 457
FT CARBOHYD 497 497
SQ SEQUENCE 520 AA; 55643 MW; 16469DC547C61F3C CRC64;

Query Match 6.9%; Score 212; DB 1; Length 520;
Best Local Similarity 21.5%; Pred. No. 5.4e-08;

Matches 119; Conservative 60; Mismatches 219; Indels 156; Gaps 23;

```
OY 34 HEKRYAQSPPYPMETVLPPIPVKQRLVTNPVNGEIMWEYELKPFHQVYDGL 93
DB 15 HSEGRVS-----AAGPVTG--LTISN-----GDVSPDGF 42
OY 94 SADLVGYDQMSGPTFOVPRGVETVVRFINNAE-----APNSVHLGSSRAAFDGMAR 147
DB 43 TAAAVLANGVPPGLITGKGNFQINVIDNLSTETMLKSTSIHMGFPQKGT--NMAAG 100
OY 148 -----DTEPGSKRDYPPYRNQASRTIYHDHAMHTAENNAKQACLYMLTPRAEDA 200
DB 101 AAFVNOCPPLATGNSFL--YDFTATDQAGTFWYHSH--LSTGYCDGLRGPMMVYDPSDPHADL 158
OY 201 LNLPSGYGFDIPMLITSKOYTANGNL-----VTTNCELNSFWMD-----VIHV 244
DB 159 YVY-----DDETTITLISWYHTAASLGAAPFGSDSTILNG--LGRFAGGSDTLAVITV 212
OY 245 NGQWPFKNVEPRKYRFRFLDAVSRSF-----GLYPADTDALIDTRLPRFYIASDSGL 297
DB 213 E-----QGRKRYRMRLSLSCDPNVYSIDGHNMTIIEADAVN-----249
OY 298 LEHPADTSLIYSMAERYVVF-----DESDYAGKTIELNLGSGIGCTDIDYDNTDKV 353
DB 250 -HEPLTVDSIQIYAGORSFVLTPADODIDNFIHALPSAGTTSFSDGINS-----AI 300
OY 354 MRFVVA---DDTPTOPDQSVSP---ANLRDVPFSPPT-----TNPDRFGRGTGPT 398
DB 301 LKYSAGSEVDPTTETTSVLPIDEANL--VPLDSPAAPDPNIGGDVALNIDFNEDGNG 358
OY 399 WTINGVAFAADVONRLANVPVGTVERMELIANGN-----GWTM 436
DB 359 FEINQVSEFSPVYLIQILISGTTASADLLPSGSLFAVPSNSTIEISFITAANAPGACH 418
OY 437 PHIHLDVDEKVSRTSGNNAFTVMPYEGSLKDYVMLGRETYVEAHYAPFGVYVFNCH 496
DB 419 PPHLGHFTSIVRTKSGSDTTFNVPVR---RDVNTGVGVGDNVITRETTDNGPMPWLHCH 475
OY 497 NLIHEDHDMMAFN 510
DB 476 IDFHLEAGFAIVFS 489

RESULT 12
LACS_TRAVE STANDARD: PRT: 527 AA.
AC 012717:
AC 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Laccase 5 precursor (EC 1.10.3.2) (Benzenediol: oxygen oxidoreductase)
DE (Urishiol oxidase) (Diphenol oxidase) (Laccase IV).
DE LACS OR LCCIV.
OS Trametes versicolor (White-rot fungus).
OC Eukaryota; Fungi; Basidiomycota; Homobasidiomycetes;
OC Aphyllophorales; Coriolaceae; Trametes.
OX NCBI_TaxID=5325;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=527;
RX MEDLINE=97464057; PubMed=9322748;
RA Ong E., Pollock W.B., Smith M.;
RT "Cloning and sequence analysis of two laccase complementary DNAs from
RL the ligninolytic basidiomycete Trametes versicolor.";
RL Gene 196:113-119(1997).
CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
CC PRODUCTS (PROBABLE).
CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzoquinone + 2
CC H(2)O.
CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
```

```
CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC DR EMBL; U44431; AAC49829.1; .
CC DR HSSP; P37064; IAO2.
CC DR InterPro; IPR001117; Cu-oxidase.
CC DR InterPro; IPR002355; Multicopper oxidase.
CC DR Pfam; PF00394; Cu-oxidase; 3.
CC DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; FALSE NEG.
CC DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; FALSE NEG.
CC KW Oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;
CC Lignin degradation; Multigene family.
CC FT SIGNAL 1 23
CC FT CHAIN 1 24
CC FT DOMAIN 25 150
CC FT DOMAIN 162 306
CC FT DOMAIN 373 498
CC FT METAL 87 87
CC FT METAL 89 89
CC FT METAL 132 132
CC FT METAL 133 134
CC FT METAL 425 425
CC FT METAL 428 428
CC FT METAL 430 430
CC FT METAL 480 480
CC FT METAL 481 481
CC FT METAL 482 482
CC FT METAL 486 486
CC FT METAL ? 74
CC FT CARBOHYD 77 74
CC FT CARBOHYD 156 156
CC FT CARBOHYD 209 209
CC FT CARBOHYD 233 233
CC FT CARBOHYD 242 242
CC FT CARBOHYD 276 276
CC FT CARBOHYD 317 317
CC FT CARBOHYD 358 358
CC FT CARBOHYD 366 366
CC FT CARBOHYD 393 393
CC FT CARBOHYD 402 402
CC FT CARBOHYD 464 464
CC SQ SEQUENCE 527 AA; 56094 MW; D9597491F1F79825 CRC64;
```

Query Match 6.8%; Score 210; DB 1; Length 527;
Best Local Similarity 22.6%; Pred. No. 7.7e-08;
Matches 119; Conservative 61; Mismatches 214; Indels 132; Gaps 26;

```
OY 88 VYPD--LGSADLVGYDQMSGPTFOVPRGVETVVRFINNAE-----APNSVHLGSSRA 140
DB 38 VPPDGTTRAVALAG--GVPPGLITGKNKDEQIIVINDLTMETMLKSTTIHMGIFQ-- 93
OY 141 AFDGMAR-----DTEPGSF-KDYPPYRNQASRTIYHDHAMHTAENNAKQACLY 191
DB 94 ACTINNADGAFAFNQCPPIATGNSFLYDFVYDQ--AGTFWYHSH-----LSTGYCDGLRPL 147
OY 192 MLTPRAEDALNLPSCYGEEDIPMLT-----SKQYANGNLVTTNCELNSFWMD 240
DB 148 VYDPPDANASL--YVDVDDTFTVITLADWYHTAALGAPAPAGPSVYLNG--LGRFSGD 203
OY 241 -----VIHNGQWPFKNVEPRKYRFRFLDAVSRSFGLYPADTQALIDTRLPRFYIA 292
DB 204 GGGATNLVITV-----TQGRKRYRRLVSLSCDPNF-----TSSIDGH---NMTI 245
```


DB 474 PGWFLCHIDFH 486

RESULT 14

LAC2_CORHI STANDARD: PRT: 520 AA.

AC 002497;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Laccase precursor (EC 1.10.3.2) (benzenediol:oxygen oxidoreductase)

DE (Urishiol oxidase) (Ligninolytic phenoloxidase).

OS Coriolus hirsutus.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Coriolaceae; Coriolus.

OX NCBI_Taxid=5327;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN-IFO 4917;

RA MEDLINE-90368706; PubMed-2394718;

RA Kojima Y., Tsukuda Y., Kawai Y., Tsukamoto A., Sugijura J.,

RA Sakano M., Kita Y.;

RT "Cloning, sequence analysis, and expression of ligninolytic phenoloxidase genes of the white-rot basidiomycete Coriolus hirsutus."

RT J. Biol. Chem. 265:15224-15230(1990).

CC -1- FUNCTION: MOST PROBABLY PLAYS AN IMPORTANT ROLE IN LIGNIN DEGRADATION. CLEAVES THE C-C AND C-O BONDS OF SOME PHENOLIC LIGNIN MODEL COMPOUNDS (SUCH AS O- AND P-QUINOLS, AMINOPHENOLS AND PHENYLENEDIAMINE).

CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) -> 4 benzoemiquinone + 2 H(2)O.

CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED BINUCLEAR.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- POLYMORPHISM: 2 ALLELIC FORMS VARYING IN ONE AA POSITION.

CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.

CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

DR EMBL: M60560; AAA33103.1; -

DR EMBL: M60561; AAA33104.1; -

DR PIR: A35883; A35883.

DR InterPro: IPR001117; Cu-oxidase.

DR InterPro: IPR002355; Multicu_oxidase2.

DR Pfam: PF00394; Cu-oxidase; 3

DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.

DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; FALSE_NEG.

KW Oxidoreductase; Signal; Copper; Metal-binding; Lignin degradation;

KW Glycoprotein; Repeat.

FT SIGNAL: 1 21

FT CHAIN: 22 520

FT DOMAIN: 23 148

FT DOMAIN: 160 302

FT DOMAIN: 369 491

FT METAL: 85 85

FT METAL: 87 87

FT METAL: 130 130

FT METAL: 132 132

FT METAL: 416 416

FT METAL: 419 419

FT METAL: 421 421

FT METAL: 473 473

FT METAL: 474 474

POTENTIAL.

LACCASE.

PLASTOCYANIN-LIKE 1.

PLASTOCYANIN-LIKE 2.

PLASTOCYANIN-LIKE 3.

COPPER (TYPE 2) (BY SIMILARITY).

COPPER (TYPE 2) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 1) (BY SIMILARITY).

COPPER (TYPE 2) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 3) (BY SIMILARITY).

COPPER (TYPE 1) (BY SIMILARITY).

FT METAL 475 475

FT METAL 479 479

FT METAL ?

FT CARBOHYD 50 50

FT CARBOHYD 72 72

FT CARBOHYD 75 75

FT CARBOHYD 210 210

FT CARBOHYD 229 229

FT CARBOHYD 354 354

FT CARBOHYD 457 457

FT DISULFID 106 509

FT DISULFID 138 226

FT VARIANT 411 411

FT CONFLICT 378 379

SQ SEQUENCE 520 AA; 55688 MW; 977D8DFA551F7929 CRC64;

Query Match 6.7%; Score 205.5; DB 1; Length 520;

Best Local Similarity 23.0%; Pred. No. 1,6e-07; Indels 137; Gaps 29;

Matches 121; Conservative 67; Mismatches 202;

QY 87 QVPPDLGSADLVGDMSPGPTFQVPGVETVVFRTINAE-----APNSVHLGSEFSRAA 141

DB 35 EVSPDGFARQAVVNVNVTGELVAGNKGDFOLNVDNLNHTMLKSTSHMGFFOKG 94

QY 142 FEGMAE-----DITEGSEF-KDYIYRNKOSARLWTHDHAMHTTAENAYRGQGLW 192

DB 95 --WMDGPAFVNOCPISSGHSFLYDFOVPOQ--AGTFWVHSH---LSTQYCDGLGPFV 146

QY 193 LTPPADALNLPGSGCEFDPMILTSKOYTA-----NG---NLVTNGE 233

DB 147 VYPPNDPHASLYVND-DVITLADMYHTAAKLGAFPLGADATLNLGGRSPSTTAAD 205

QY 234 LNSFWGDVHVNGQWPFKNVEPRKYRFRFLDAAVSRSGFLPADYDAIDFLRPFVIAS 293

DB 206 L-----AVINV-----TKGRYRFRVSL-----SCDPNHTFSIDGH 237

QY 294 DSGLL-----HPDTSLLYISMAREYEVFPFSDYACKTIELR---NLG--GSIGGIC 342

DB 238 DLTIIEVDSINSQPLVVDSTQIFPAQRYSPVLNADODVG--NWMIRANPNFGNVGAGIN 296

QY 343 TD-TDVNDMDKVRVFAVADDTQPDTSVVPANLRD-----VPPSP---TNTPRGF 390

DB 297 SALLRTDGDAPV-----EPTTQTTPTRKPLNEVDLHLPLATAVAP--GSPVAGVDIAIM 349

QY 391 RFSRTGPTWTINGVAPADYONRLLANVPYGVYERWELINAGN----- 432

DB 350 AFNFGNTPFINGASFPVPPVPLQLIISGAQNAQDLPSGSVSLPSNADIEISPPATA 409

QY 433 ---GTHPHIHLVDFKVISRTSGNNARVMPYESGL-KDYVNLG---RRETVVEANHA 485

DB 410 AARGAPHPFLHGHAFVAV--RSAGS---TVYNNDNIFRDVYSTGTTPAAGDNTIRFR--T 464

QY 486 PRFGVYMFHCHNLIHEDDHMMAFNAFVLPDYGNATVFPVDPMEELM 532

DB 465 DNGPWFHLCHIDFHLEAGFAVVF-AEDIPD-----VASANFPQAM 505

RESULT 15

LAC2_TRAVE STANDARD: PRT: 519 AA.

AC 012718; 012718;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Laccase 2 precursor (EC 1.10.3.2) (Benzenediol:oxygen oxidoreductase)

DE (Urishiol oxidase) (Diphenol oxidase) (Laccase I).

GN LCC2 OR LCC1.

OS Trimetes versicolor (White-rot fungus).

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Coriolaceae; Trimetes.

OX NCBI_Taxid=5325;

RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-52;
 RX MEDLINE-97464057; PubMed-9322748;
 RA Ong E., Pollock W.B., Smith M.;
 RT "Cloning and sequence analysis of two laccase complementary DNAs from
 RL the lichenolitic basidiomycete *Trametes versicolor*.";
 CC Gene 196:113-119(1997).
 CC -1- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
 CC PRODUCTS (PROBABLE).
 CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzoquinone + 2
 CC H(2)O.
 CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
 CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
 CC 3 OR COUPLED BINOCULAR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.
 CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isd-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U44851; AAA86659.1; -
 CC EMBL: U44430; AAC9828.1; -
 CC HSSP: P37064; IAOZ.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR002355; MultiCu_oxidase2.
 DR Pfam: PF00394; Cu-oxidase; 3.
 DR PROSITE: PS00079; MULTICOPPER OXIDASE1; 1.
 DR PROSITE: PS00080; MULTICOPPER OXIDASE2; FALSE_NEG.
 KW Oxidoreductase; Signal; Copper; Metal-binding; Glycoprotein; Repeat;
 KW Lignin degradation; Multigene family.
 KM SIGNAL 1 20
 FT CHAIN 21 519
 FT DOMAIN 22 147
 FT DOMAIN 159 301
 FT DOMAIN 368 490
 FT DOMAIN 84 84
 FT METAL 86 86
 FT METAL 129 129
 FT METAL 131 131
 FT METAL 418 418
 FT METAL 420 420
 FT METAL 472 472
 FT METAL 473 473
 FT METAL 474 474
 FT METAL 478 478
 FT METAL 74 74
 FT CARBOHYD 161 161
 FT CARBOHYD 228 228
 FT CARBOHYD 237 237
 FT CARBOHYD 271 271
 FT CARBOHYD 353 353
 FT CARBOHYD 361 361
 FT CARBOHYD 456 456
 FT CONFLICT 69 69
 SQ SEQUENCE 519 AA; 55810 MW; BFFB5B4CD0007702 CRC64;

Query Match 6.7%; Score 205; DB 1; Length 519;
 Best Local Similarity 23.7%; Pred. No. 1.7e-07;
 Matches 124; Conservative 63; Mismatches 203; Indels 134; Gaps 30;

OY 88 VYPDGSADLVGCTDMSRPTQVRGVETVVRFINNAE-----APNSVHLGSPS 138
 DB 35 VSPDGLRDAIVNGVSPSLITGRKG---DRFOLNVDDTLNHSMLKSTSIHWHGFPQ 90

OY 139 RAAFDGMAE-----DITEPGSF-KDYVYYPNROSARTLWYHDHAMHITAENAYRGOAG 189
 DB 91 -AGTNWADGPAFVAVNOCPLASGHSFLYDFHPVDO--AGFTWYHSH-----LSTQYCCGLNG 142
 OY 190 LYMLTDPADDAALNPSGYSGBEDIP-----MILTSKOYTANGNLVTNGELNSFWGDV 244
 DB 143 PFVYVYDPKDPHAS-----RYVDNNESTVITLTDWYHTA-----ARLGRFPLGADATL 191
 OY 245 NG-----QPMPEKANYE-PAKYRFRFLDAVNSFGILYFADTADTRLPFKYIAD 294
 DB 192 NGIGRSASTPTAALAVINVOHGKRYRFLVSIISCDPNY-----TFSDIGH-NLFTVIEVD 244
 OY 295 SGLLEPADTSLIYSMAERYEVYFDFSDYAGK--TIELRNIG--GSIGIGTDTDYDMT 350
 DB 245 -GINSQPLVVDSTQITPAOQTSFYVLNANOTGVNGYWRANPNSTGVFAGINS----- 296
 OY 351 DKVMRF--VVADDTTQPTSVVP--ANLRDVP-PPSPPTNTP-----RQFRFGRTGP 397
 DB 297 -ALIRYQAPVAPVPTTQTSVYPLIETMLHPLARMVPGSPTPGGVDKALNLAFFNGT 355
 OY 398 TWITNGVADADVONRLLANVPVGVTVRWEELIANGN-----GWTN 436
 DB 356 NEFINNATPTPTVPVLLQILSGAQTAODLLPAGSVYPLPAHSTIETPLPATYALAGAP 415
 OY 437 PIHILVDEKVISRTSGNNARFVMPYESGL-KDVVWLG--RRETVVEAHVAPFGGYV 492
 DB 416 PRHLGHAFAYV-RSAGS---TTYNTNDPIFRDVSSTGTPAAGDNTIR-FQTDNFGPMF 470
 OY 493 FPHCNLIHEDHDMMAAFNATVLPDGYNATVFPVDPMEELMQARP 536
 DB 471 LRCHIDPFLD---AGF-----ALVFAEDVADYKAAMP 499

Search completed: October 3, 2002, 11:02:18
 Job time: 442 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 11:03:21 ; Search time 55.98 seconds

(without alignments)
1767.651 Million cell updates/sec

Title: US-09-656-640A-4

Perfect score: 3067

Sequence: 1 MFKHTLGAALSLIFNSNAV.....AVTERIQTMAREYRAADE 572

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteriap:*
17: sp._archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1913	62.4	602	3	Q9P8C3
2	653.5	21.3	475	2	Q93M03
3	506	16.5	568	10	Q23123
4	502	16.4	637	10	Q9FTS3
5	495	16.1	582	10	Q9AMU4
6	485	15.8	581	10	Q949Y9
7	470	15.3	591	10	Q9C9A4
8	456	14.9	591	10	Q9FTS5
9	454.5	14.8	614	10	Q9FTS6
10	400.5	13.1	527	16	Q67206
11	400	13.0	536	2	Q938E6
12	393	12.8	515	16	Q9CJ06
13	384.5	12.5	502	16	Q98F88
14	288	9.4	468	16	Q9CPE1
15	276.5	9.0	1662	2	P71431
16	276	9.0	513	16	Q9PME8

17	266.5	8.7	460	2	Q9X3V2	Q9X3V2 pseudomonas
18	261	8.5	463	16	Q9HXW7	Q9HXW7 pseudomonas
19	260	8.5	500	16	Q9K8A9	Q9K8A9 bacillus ha
20	242.5	7.9	520	3	Q95944	Q95944 ceriporiops
21	241.5	7.9	504	16	Q53858	Q53858 mycobacteri
22	229.5	7.5	518	3	Q59896	Q59896 pycnoporus
23	229	7.5	464	16	Q92QV6	Q92QV6 rhizobium m
24	225.5	7.4	518	3	Q9UVO2	Q9UVO2 pycnoporus
25	225.5	7.4	561	3	Q96WNO	Q96WNO dolitylus cl
26	224.5	7.3	520	3	Q13448	Q13448 coriolus ve
27	223.5	7.3	520	3	Q96U77	Q96U77 trameles ve
28	222.5	7.3	518	3	Q9HDS9	Q9HDS9 polyporus c
29	221.5	7.2	518	3	Q96RR6	Q96RR6 pycnoporus
30	220.5	7.2	518	3	Q96V45	Q96V45 pycnoporus
31	220	7.2	539	3	Q9Y780	Q9Y780 coprinus cl
32	220	7.2	567	10	Q941X2	Q941X2 oryza sativ
33	220	7.2	569	10	Q9FY79	Q9FY79 arabidopsis
34	217.5	7.1	519	3	Q13420	Q13420 basidiomyc
35	216	7.0	520	3	Q93F57	Q93F57 pseudomonas
36	215.5	7.0	520	3	Q94222	Q94222 trameles ve
37	215.5	7.0	524	3	Q13422	Q13422 basidiomyc
38	215.5	7.0	589	3	Q9C497	Q9C497 glomerella
39	215	7.0	352	2	Q93F48	Q93F48 pseudomonas
40	215	7.0	586	10	Q9AU16	Q9AU16 pinus taeda
41	214.5	7.0	351	2	Q93F47	Q93F47 pseudomonas
42	214.5	7.0	526	3	Q13456	Q13456 coriolus ve
43	214.5	7.0	577	10	Q9FUD5	Q9FUD5 arabidopsis
44	213.5	7.0	353	2	Q93F51	Q93F51 pseudomonas
45	213.5	7.0	533	3	Q60199	Q60199 pleurotus o

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	602 AA.
Q9P8C3	Q9P8C3			
AC	Q9P8C3:			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	POLYPHENOL OXIDASE PRECURSOR.			
GN	PPOA.			
OS	Acromonium murorum.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;			
OC	Hypocreales; Hypocreaceae; mitosporic Hypocreaceae; Acromonium.			
OX	NCBI_TaxID=45278;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CBS 157.72;			
RX	MEDLINE=21268855; PubMed=11375170;			
RA	Gouka R.J., van der Heiden M., Swarthoff T., Verrips C.T.;			
RT	"Cloning of a phenol oxidase gene from Acromonium murorum and its			
RT	expression in Aspergillus awamori."			
RL	Appl. Environ. Microbiol. 67:2610-2616(2001).			
DR	EMBL: AJ271104; CAB75422.1; -			
KM	Signal.			
FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	62	602	POLYPHENOL OXIDASE.
SQ	SEQUENCE	602 AA; 66920 MW; 0303D991405228A3 CRC64;		
QY	Query Match	62.4%; Score 1913; DB 3; Length 602;		
QY	Best Local Similarity	61.0%; Pred. No. 2.7e-135;		
QY	Matches 365; Conservative 73; Mismatches 126; Indels 34; Gaps 8;			
DB	5 TLGAALSLIFNSNAVASP-----VPE-----TSPAGHLFKRY 39			
DB	4 TTRLRALALYLSLKGAAQAMKFEIDPEEAAALAAIVEDDPANDIQRRSPAD--LQRR 61			
QY	40 AQTSPQPMETVLPPIPKOPRLTVTNPGQEIYVEVEIKPFTHQVYPDLSADLVG 99			
DB	62 SPLSPAVYTLFGQAPLSTIPVKEPLFTVTNPNNGCEIDYVEIEIKHFSQVFPDLCPADLVG 121			

Oy	100	YKDSPGCTEQVPRGVEVYVFFINNAEPNSVHLHGSSRAFPQGMEDITEPESGDY	129
Oy	122	YDGISPGCTEQEKGRESVVFFVKAIVYESSIHLHGSSRAFPQGMEDVITNGEYVDY	161
Oy	160	YPNROSARTLWYHDBAHMHTAENARYGQAGLMLTDPDAEDALNDPSGGEFIDPMLTSK	219
Oy	182	YPNAGAGRSWYHDBAHMHTAENARYGQAGLMLTDPDAEDALNDPSGGEFIDPMLTSK	241
Oy	220	OYTANGNLVTTNGELANSWGVIVHYNQGPWFVKVVERPKTRFRFLDAVAARSGLYFADT	279
Oy	242	QYNSDGLTFESTKGEQSLMGVIVONGVPMWPEFVEVERPKYRFRFLDAVAARSFSLYVDT	301
Oy	280	DAIDRLPEFKYIASDGLLEHPADTSLIYIMAREYEVGFSPDQACKTIELRLNGSIG	339
Oy	302	ADEDTIRIFQVIYASDGLLEEPYTTSKLYSIARREYELILDFSPEKKTIELRN-EPAGV	360
Oy	340	GIGTDTDYDNTDKWRFVYVAD-DTTPDTSVYVPAKLRDVPFPPPTNT-PROFAGRTGP	397
Oy	361	GLGIEVNYDDDTDKWRFVYVAGLSPSPDTSVYVPTSLRDVPFPPSTITDHSFRTAG	420
Oy	398	TWTINGVAFADVONRLANVGVGVVERKLELINAGNGTTHPIHLVLPFKYISRI--SGN	454
Oy	421	QMSINGVGFSDVENKLLANVPLGTVOIMQLTNAAQGTTHPIHLVLPFKILSRGAATGA	480
Oy	455	NARTWMPES-GLKDYVWLGRETVYVEAHYAPRPGVYMFCHNLIHEDHDMMAAFATV	513
Oy	481	TTTQVEPEPSAGLNDVYILGGEIVYVYEAHAPRPGVYMFCHNLIHEDHDMMAAFATV	540
Oy	514	LPDGYNATVADVPMEELMQRAPYELGEFOAQSGFSVQAVTERIQTMAEYRPPYAAD	571
Oy	541	LPDGYNSTALADPQDEFFRAKPYVDNDVEVENAFSTDEIAEQVLMASLYLPYDND	598
RESULT	2		
Q93M03		PRELIMINARY;	PRT; 475 AA.
AC	O93M03;		
DT	01-DEC-2001 (TREMblrel, 19, Created)		
DT	01-DEC-2001 (TREMblrel, 19, Last sequence update)		
DT	01-DEC-2001 (TREMblrel, 19, Last annotation update)		
DE	OXIDOREDUCTASE-LIKE PROTEIN.		
GN	AUR21.		
OS	Streptomyces aureofaciens.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;		
OC	Actinomycetales; Streptomyicinaeae; Streptomycetaceae; Streptomyces.		
OX	NCBI_TaxID=1894;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CJM3239;		
RA	Korhaneec J., Bistakova J., Novakova R., Homeroova D., Rezuchova B.,		
RT	"Cloning and characterization of a new polyketide gene cluster in		
RL	Streptomyces aureofaciens CJM3239."		
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY03994; AAK61713.1;		
DR	SEQUENCE 475 AA; 52602 MW; A2ED9A4C638DDOC9 CRC64;		

[illegible]

```

0Y 219 KQYANGULVJTTNGELNSFMGDVJH--VNGQPMPEFKVNEPKRYEFLDLAAVSSELYE 276
Db 200 AVEVEDGTLTLTRP-----SDCEMLVNGKEREYFOVAARKERIRYNNACANNKYKLR 253
0Y 277 ADTDAIDTRLEPFKVIASDGLLEHPADTSLIYISMAERYEVPEDFSOY-AGKTIELNIG 335
Db 254 AD-----GIEFTQGTGGFLERPVQOSELMLGGERADLVNDFSRRKXGDSVLENP 307
0Y 336 G-SIGIGTDTDYNDTKVMRFVVAADDTTOPDTSVPAHLKDPFSPSTNTPROFRGR 394
Db 308 AGSI-----ERPEVMRFELV--RTADYSFVFGRLTYV-PQPTLYVERDFEI-R 353
0Y 395 TGEPTTNGVAFADYQNRLLANPEYGERKELIN---AGNG-----WTHPIHILVD 444
Db 354 TEPAPMTINGOSYD--PNEVDITAKLTGTEVVTVAENVEAPARCAKRPDEFLM--HSEPHILTY 410
0Y 445 EKVIERTSGNNAKRYMPEESGLKDVWLGRETVVVEAHNAPFGVGMFCHNLIHEDHD 504
Db 411 FVFLER----NGKPAGRDLGLKDTVLTLLGPEDIVKIAMTWGPRYGOYLHYCHQHGSSGG 466
0Y 505 MM 506
Db 467 QM 468

```

RESULT	3	
ID	023123	PRELIMINARY; PRT; 568 AA.
AC	023123.	
DT	01-JAN-1998 (TREMBLrel. 05, Created)	
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE	F19G10.5 PROTEIN.	
GN	F19G10.5.	
OS	Arabidopsis thaliana (Mouse-ear cress).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spemathophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsis.	
OX	NCBI_TaxID=3702;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CV. COLUMBIA;	
RA	Federstepfel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,	
RA	Au M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,	
RA	Oji O., Osborne B.I., Shinn P., Sun H., Toriumi M., Wysockaia V.,	
RA	Yu G., Ecker J., Theologis A., Davis R.W.;	
RL	Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: AF000657; AAB72167.1; -	
DR	InterPro: IPR001117; Cu-oxidase.	
DR	Pfam: PF00394; Cu-oxidase; 1.	
SO	SEQUENCE 568 AA; 64355 MW; C4314C889576E35E CRC64;	

[illegible]

```

OY 229 ---GDVIVHGQGWPEKKNVEPRKRYRFFELDLAAVSRSGLEFADDDAIDTRLPEKVIAS 293
      |||||
DB 254 OPEYFGDIIYVNGKAMPRLNRRRYRREFTIINASNAKRFEPFSN-----GIDPLVGS 307
      |||||
OY 294 DSGLEHREADTSLIATISMAERYEVVEFDESDYAGKTIELNLGSGIGIGIDPDY----- 347
      |||||
DB 308 DSAVLSKRVPMRKSLISLSESLVDVVVDFKSPSRIVYLAN-----DAPYPYPSGD 357
      |||||
OY 348 ---DNTDKVMRNVVADDTTQPTSVYVPAVLKRVYPPSPPTTNPROFRFRTGPTWINGV 404
      |||||
DB 358 PVNENGVGMKFFII--NNSESDTCTIPKKL--INYNADVSAVNLRY-----ISMYEV 409
      |||||
OY 405 AFADQNQKLLAN-----VPV-----GTYERMELINAGGWHPIHILVDEKVISRTS- 452
      |||||
DB 410 SNSDEPTHLVNGLEPYEARVYTPETKSGTTEVWEVNLTED--NHPLHILGLFRVYEQTLV 468
      |||||
OY 453 -----GNNA-----RTVPEYSGLDVYVL--GRRETVVEV 483
      |||||
DB 469 LAAGLEEFKECMTKQNDAAVKQISKYAROKTKAYVAHAHEGKMNVFMFMGHYTRILVRPS 528
      |||||
OY 484 YA-----PF-----PCYYMFHCHLILHEDHDM 506
      |||||
DB 529 YHTNASYPFDPTQBPQ--YVYHCHILHEDHNNMM 561
      |||||

RESULT 4
O9FTS3 PRELIMINARY; PRT; 637 AA.
ID O9FTS3
AC O9FTS3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE SPORE COAT PROTEIN-LIKE PROTEIN.
GN P0409B08.14.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eubartoideae; Oryzaceae; Oryza.
RX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
RT clone,P0409B08."
RL submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002860; BABI8287.1; -
DR InterPro; IPRO02355; MultiCu_oxidse2.
DR PROSITE; PS00080; MGLTICPPER_OXIDAS52; UNKNOWN_1.
KW Coat protein.
SQ
SEQUENCE 637 AA: 70919 MW: 652D39F10F2D6A96 CRC64:

```

Query Match	16.4%	Score 502	DB 10	Length 637
Best Local Similarity	29.1%	Pred. No. 2	56-29	
Matches	174	Conservative	72	Mismatches 186; Indels 164; Gaps 29
OY	24	PVPETSPATG-----HLEK-----RVAQISPOYMETVPLRDPVKORPLTVN	67	
Db	82	PLP--APAAGRMRMVMNLSLNTLKRYVDSIRIAKIRGYGIRGRAPVI-----KTY--	131	
OY	68	PVNG--OELIYWEVEIKRPTHQVYRDLSADLVGDGM-----PGPTEQYQGVETV	119	
Db	132	---GMYSKTQWPHRMDPRFPYVY-----GCSLQATPRPGPTIYARQGVHLAV	176	
OY	120	RFINN-----AEAPNSVHLGSEFRAAFDMA-----EDLIEPS	154	
Db	177	EQMNIHLPAHLIPMDPKVPTAIRKKGGVPTVHLHGAPREFDGHAFAMFTRDPFENS	236	
OY	155	F---KDYUUPNQASRLTYNDHANHITAEKNVYRQAGLYMLTDP--AEALNLSSTGE	210	
Db	237	TWTKRTYTPPNQAGCNITLNDHAGLITFVSLAGLILAVIYEKRELEDPKMLPC--GGH	294	

[illegible]

	RESULT	5	
ID	09AMU4	PRELIMINARY:	PRT: 582 AA.
AC	09AMU4;		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DE	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
GN	PURATIVE SPORE COAT PROTEIN.		
OS	Oryza sativa (Rice).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OX	Ehrhartioideae; Oryzeae; Oryza.		
NCBI_TaxID=4530;			
[1]			
RN	SEQUENCE FROM N.A.		
RC	STRAIN-CV. NIPPONBARE;		
RA	Sasaki T., Matsumoto T., Yamamoto K.;		
RT	"Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC		
RL	clone:P0044F08.";		
RP	Submitted (NOV-2000) to the EMBL/genbank/DBJ databases.		
[2]			
RN	SEQUENCE FROM N.A.		
RC	STRAIN-CV. NIPPONBARE;		
RA	Sasaki T., Matsumoto T., Yamamoto K.;		
RT	"Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC		
RL	clone:P0037C04.";		
RP	Submitted (FEB-2001) to the EMBL/genbank/DBJ databases.		
DR	EMBL: AP002909; BAB21188.1; -		
DR	EMBL: AP003233; BAB55542.1; -		
DR	InferPro; IPRO02355; MultiCu_Oxidase2.		
DR	PROSITE: PS00080; MULTICOOPER_OXIDASE2; UNKNOWN_1.		
KW	Coat protein.		
SQ	SEQUENCE 582 AA; 64050 MW; B5054EECB2C5CIDA CRC64; -		
Query Match	16.1%; Score 495; DB 10; Length 582;		
Best Local Similarity	29.28%; Pred. No. 7.6e-29;		
Matches 175; Conservative	73; Mismatches 219; Indels 132; Gaps		
0Y	9 AALSLIFNSNAQAOSPVPETSPATGHLERKVA-----QISPOXYM---FTV---PLP 54		
Dd	8 AVLLLVIGVAAGCTRRPSAPRPVTEDTLQKVAGSLEMYVDLPQRKIVGCFGRHGHSP 67		
0Y	55 IPPVQPRILTYINPNVG--QELTWYEVELEKPTTHQYVRDLSGADLVGYGMSRGPTFOVP 112		
Dd	68 I-----RLTI-----GMIOKKMKRHRRDPASTVEFGTSAAT-----ATPPGPITEAA 110		

Qy	113	RGELVYAFINMAEA-----PMSYHLHG-----SFSRAAFDSMAE	147
Db	111	QGVPLSTVMQWYLEARHRLHMPDTPVTPAIPRRGGVPTVYVHLHGAAHPQSGDSAFAMWETA	170
Qy	148	DITEFG---SFKDYVYPPRQSAFTLAWHDHAMHITAENAAVGGAGLYMLDPAEDA ¹ LA ¹	203
Db	171	GGEGFGAWMSFTPTTYTPAQSQVYLMWHDHALGLTRANLLAGLLGAVINNPVAVELGL	230
Qy	204	PSGGEEDIPILTSKOYTANGNL-VTTNGEFLN-----EMGDVYHVGQWPEPKNV	254
Db	221	PCG-DEFPRVLMILDRSPYADGSIYMNATGIIPIHPOWQPEYGEAITYNGKAMPLAV	289
Qy	255	EPRKRFPEFLAAVSRSGLYEAFDIDAIDTLPFKVIAASDGLLEHPADSLIYISMAER	314
Db	290	ARRRFRFIITNSARRFNL-----SITNGLPFTVVGSDTNYLSKRYTASLLVSAET	343
Qy	315	YEYVDFESDYAGKTIELRNLGSGIGIGTDTDYDNTDKVMKFEVVAADDTTOPDTSVPANL	374
Db	344	FDVVVDFQSSTSEAEVNTAPFPYPDG-QAPNDLNGKVMKFEVISPAAK-DTSRPVAKL	401
Qy	375	RDVEPPTSTTTPR-----QFRGRG-PTWT-INGVAPADVQNRLLANPVGYVERWE	426
Db	402	LDYVAVAEEAEVQRYIYWEYEDAAATGNPHILYXNGKRLDEPATE--TPRQTEYWE	458
Qy	427	LINAGNGWTHPIHILHLDVFKV-----ISRTSGNNATVM	460
Db	459	VINLTPD-NHPLHLHLATFOATRGVLGVDDEAEFKGCAKMLNDVRCVSRHAAVEEAVP	517
Qy	461	PYESGLKDVYML--GRRETVVEAHYA-----PF-----PGYVMFCHNLIHEDHDM	506
Db	518	EHEKGWKNVYVIAAGYMTTIVYKFEFMYDSGKRPFDATAEFG-VYVYCHILHDEDMNI	575

RESULT 6

Q949X9 PRELIMINARY; PRT; 561 AA.

Q949X9

AC 01-DEC-2001 (TREMblrel, 19, Created)

DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)

DT 01-DEC-2001 (TREMblrel, 19, Last annotation update)

DE PUNITATIVE SPORE COAT PROTEIN.

GN F23N20.3.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eumariophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucosids II; Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxId=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,

RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,

RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,

RA Carlsbaci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,

RA Kamitani A., Kerlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B.,

RA Lin J., Meyers R.C., Miranda M., Natuska M., Nguyen M., Palm C.J.,

RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Tracy S.R.,

RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,

RT "Full length cDNA of gene F23N20.3 (GI:12323429)."

RL submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL, AY050818; AAK92753.1; .

QO Coat protein.

SEQUENCE 581 AA; 66165 MW; 64873EA3FEC24FBD CRC64;

Query Match	15.8%	Score 485	DB 10	Length 581
Best Local Similarity	30.2%	Pred. No. 4.2e-28		
Matches 156; Conservative	66;	Mismatches 183;	Indels 112;	Gaps 23

```

0Y      91 DLGSADLVGYDCMS-----PGPTFQVPRGVETVVRFIN-----N 124
      || : : || |      |||| : || : | : |
Db      84 DLPRATPVFAV-GTSKRSATVGPPTTEAVGVVDVYVTVWRNHLPLNHLPLWDPPTISPAIPKH 142

```

[illegible]

RESULT	7		
09C9AA		PRELIMINARY;	PRT; 591 AA.
ID	09C9AA		
AC	09C9AA;		
DT	01-JUN-2001 (TREMblrel, 17, Created)		
DT	01-JUN-2001 (TREMblrel, 17, Last sequence update)		
DT	01-JUN-2001 (TREMblrel, 17, Last annotation update)		
DE	SPORE COAT PROTEIN-LIKE PROTEIN.		
GN	F3JN20.3.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.		
NC	NCBI_TaxID=3702;		
LN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV. COLUMBIA.		
RX	MEDLINE=21016719; PubMed=11130712;		
RA	Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,		
RA	White O., Alonso J., Altai H., Arujo R., Bowman C.L., Brooks S.Y.,		
RA	Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,		
RA	Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,		
RA	Dunn P., Etgu P., Feldhlym T.V., Feng J.-D., Fong B., Fujii C.Y.,		
RA	Gill J.E., Goldsmith A.D., Johns S., Hansen N.F., Hughes B., Huizar L.		
RA	Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,		
RA	Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A.Y., Lam B.,		
RA	Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,		
RA	Liu X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,		
RA	Miltscher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,		
RA	Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,		
RA	Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,		
RA	Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,		
RA	Utterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,		
RA	Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,		
RT	"Sequence and analysis of chromosome 1 of the plant Arabidopsis		
RT	thaliana.";		
RL	Nature 408:816-820(2000).		
DR	EMBL; AC016972; AAG51692.1; .		
SW	COAT PROTEIN. 591 AA; 67358 MW; B0CE7B0E8F055AF CRC64;		

Query Match 15.3%; Score 470; DB 10; Length 591;

Best Local Similarity 29.6%; Pred. No. 5.8e-27;

Matches 156; Conservative 66; Mismatches 183; Indels 122; Gaps 24;

```
OY 91 DLGSADLVGYDGS-----PGPTFOVPRGVETVVRFIN-----N 124
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 84 DLPAATPVFAV-GTSKRSATVPGPTIEAVGVGDYVYWRNHLPLHLIPMDPTSPAIPKH 142
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 125 ALEPNVHLHGFSRRAFGMAED-----ITEGSG---FKDYYPKRSKRTLYTHDHAM 176
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 143 GGIPYVHLHGHIHEPTSGDNADSWFTAGFKETGSKTKTHYVKQOPGNMYYHDHAA 202
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 177 HITAEAAVYGOAGLYMLT-DPAEDALNPSGVEFDIPMLITSKOYTANG---NLVTN 231
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 203 GLTRVLLAGLGLSYILRRSSVESPRLPTG-REFDRPLVIFPRSRKDSITMANATGN 261
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 232 GELNSFW-----GDVIVHNGQPMPEKRVPRKRYRFRFLDAAVSRSGLYFADTDAIDTRL 286
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 262 PTHPQMOPEYEGDALIVNGKAMPRLTVRRKRYFRITNASNARFPRPFESN-----GL 315
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 287 PRVIVASDGLLEHPADTSLIYSMAERIEVDESDYAGKTITELNGLSGIGTDD 346
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 316 DFLVGSDSAYLAKPVSTSVLLAPSEIVDVLDFSKSTKAILANMAPYPYSGDPVT 375
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 347 YDNTDKVMPFVADPTDPTDSVPANLRDVPSPPTTTPROFR-----GRTGPT 398
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 376 EENS-KVMKFIIT-NKSEVDTSIIPKKL-IEYPAHVSTSTTRIIAMFEVSSIDEPT 431
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 399 WT-INGVAEADVONRLANVP-VGIVE-----RMELINAGNGWTHPIHLVDER 446
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 432 HLYINGLPY---NAVTEDEPKIGTSEFESLMITMKWMEVILNLD-NHPLHLHGLK 486
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 447 V-----ISRTSGNNARTVMPESIGKDVWL--GRETEY 478
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 487 VLEQTALVKSSEFIECMTRKNDVAKCEISYTAGNKTAATVHERGKVMFKMPGHVTXI 546
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 479 VVEAHY-----APFGVYMFCHNLJHEDHMAAFNATVL 514
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 547 LVRFSTIHSNESTSPDATOEPG-YVTHCHLHLEDHMMMRPF-AMVL 591
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 8
O9FTS5 PRELIMINARY: PRT: 614 AA.

AC O9FTS5: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)
GN PUTATIVE SPORE COAT PROTEIN-LIKE PROTEIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone: P0409B08.";
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP002860; BAB18285.1;
DR InterPro; IPR002355; MultiC_oxidase2.
DR PROSITE; PS00080; MULTICOPEPR_OXIDASE2; UNKNOWN_1.
KM Coat protein.
SO SEQUENCE 614 AA; 68321 MW; 80C6FC856E305193 CRC64;

Query Match 14.9%; Score 456; DB 10; Length 614;
Best Local Similarity 27.4%; Pred. No. 6.9e-26;

Matches 170; Conservative 76; Mismatches 205; Indels 170; Gaps 30;

```
OY 23 SPVETSPATGHLFKRAOISPOYMEFTVLPPIPVKOPRL-----VTNPVNG 71
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 36 TPADSPSPRARLRSTADL-----PKYVPLP-----QMARIGYGINOGCFPTTLTG 87
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 72 --OEIYVEVEIKPPTHNOVYPDIGSADLVGYDGS-----PGPTFOVPRGVETVVRFIN 123
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 88 MYKTKQGFHDMDEPTVEFVY-----GQSLQTATFPGPTIVARVNPVLYTWEN 135
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 124 -----NAEAPNSVHLHGFSRAADGMA-----EDITPGSF--- 155
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 136 HLPEADHILPMDPTVPTAIPKNGVPTVYHLHGAQAAPDSGHAFAWTRFPAFGSTWTQ 195
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 156 KDYYVNRQ-SARTLWYHDHMHITAEENYRGOAGLYMLTDP-AEDALNPSGVEFDIP 213
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 196 KTYTYRNVQPAQNIYTHDALGLTRASLAGLLAAYIVEMPELEMPNLPS--GEFDLH 253
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 214 MILTSKOYTANGNL-VTNGELNS-----FWGDVIVHNGQPMPEKRVPRKRYRFRRL 264
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 254 LVTDARKFNVDGTIFMDTVGAVPSVHPQMOPEYFGEVITVNGKAMPFOVQRRRRYRLRI 313
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 265 DAAVSRSGLYFADTDAIDTRLPEFKVIASDGLLEHPADTSLIYSMAERIEVDFPS-- 322
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 314 NASNARYLNIREFSN-----GLPFTVIASDATYLSRPVTSNLLSPAEIIFDIYVDFSIY 367
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 323 -DYAGKTIELRLNGSGIGIGTDTDNDKVMRF-----VYADD-----TTQPD 366
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 368 VFNATDIELNLSAPYPTPTGPANATLDGKVAEFAVSAKMOGDMPMOEPENSTIVPE 427
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 367 TSVVPANLRDVP-----FPSPPTTTPROFRGRGTPTWT--INGVAEADVONRL 414
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 428 IGVPEFAKVALAPPTMKTRYIVLYENMTSNDP-----NKTATNKLXINGRLDEDP-- 477
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 415 ANVP-GYVERMELINAGNGWTHPIHLVDERKYT----- 448
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 478 TETPIGISTELMHVILNLPD-NHPLHLAEFOAVOMLOLVDPDTFKSCMLKHNDTFACN 536
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 449 -SRTSGNNARTVMPESIGKDVWL--GRETEYVVEAHYA-----PF-----PGVYMF 494
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 537 LDOHAVALQPYPEEKRTWNVKIPPAVYTSVAFLRVLHNNMPYPFATAAG-YVH 595
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 495 CHNLIHEDHMAAFNATVLP 515
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 596 CHLHEDHDMAMIRPL--TLLP 614
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 9
O9FTS6 PRELIMINARY: PRT: 588 AA.

AC O9FTS6: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)
GN PUTATIVE SPORE COAT PROTEIN-LIKE PROTEIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone: P0409B08.";
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP002860; BAB18285.1;
DR InterPro; IPR002355; MultiC_oxidase2.
DR PROSITE; PS00080; MULTICOPEPR_OXIDASE2; UNKNOWN_1.
KM Coat protein.
SO SEQUENCE 614 AA; 68321 MW; 80C6FC856E305193 CRC64;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 10:54:52 ; Search time 25.3 Seconds
(without alignments)
552.231 Million cell updates/sec

Title: US-09-656-640A-4
Perfect score: 3067
Sequence: 1 MFKHTLGAALSLFLNSNAV.....AVTERIQTMAYRPAAD 572

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfills1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3067	100.0	572	4	US-09-401-476-4
2	2010	65.5	583	4	US-09-401-476-2
3	228.5	7.5	499	2	US-09-032-315-2
4	228.5	7.5	499	2	US-08-993-318A-2
5	228.5	7.5	499	4	US-09-399-886-2
6	228.5	7.5	499	4	US-09-396-260-2
7	228.5	7.5	499	4	US-09-576-281-2
8	228.5	7.5	520	1	US-08-462-484-2
9	228.5	7.5	520	1	US-08-441-147-2
10	228.5	7.5	520	5	PCT-US95-07536-2
11	226	7.4	620	1	US-08-940-661A-2
12	226	7.4	620	2	US-09-083-485-2
13	226	7.4	620	2	US-08-939-218A-2
14	223.5	7.3	511	1	US-08-462-484-8
15	223.5	7.3	511	1	US-08-441-147-8
16	223.5	7.3	511	5	PCT-US95-07536-8
17	220	7.2	539	2	US-09-032-315-1
18	220	7.2	539	2	US-08-993-318A-1
19	220	7.2	539	3	US-08-689-421-27
20	220	7.2	539	4	US-09-399-886-1
21	220	7.2	539	4	US-09-396-260-1
22	220	7.2	539	4	US-09-389-528-27
23	220	7.2	539	4	US-09-181-827A-27
24	220	7.2	539	4	US-09-576-281-1
25	215.5	7.0	616	2	US-08-991-531-2
26	215.5	7.0	616	2	US-09-032-315-9
27	215.5	7.0	616	2	US-08-993-318A-9

28	215.5	7.0	616	3	US-09-028-887-2	Sequence 2, App1
29	215.5	7.0	616	4	US-09-399-886-9	Sequence 9, App1
30	215.5	7.0	616	4	US-09-396-260-9	Sequence 9, App1
31	215.5	7.0	616	4	US-09-518-901-2	Sequence 2, App1
32	215.5	7.0	616	4	US-09-576-281-9	Sequence 9, App1
33	214.5	7.0	616	1	US-08-749-882A-2	Sequence 2, App1
34	214.5	7.0	616	2	US-08-539-134-2	Sequence 2, App1
35	212.5	6.9	616	5	PCT-US95-06816-2	Sequence 2, App1
36	211.5	6.9	516	3	US-08-689-421-29	Sequence 29, App1
37	211.5	6.9	516	4	US-09-389-528-29	Sequence 29, App1
38	211.5	6.9	516	4	US-09-181-827A-29	Sequence 29, App1
39	210.5	6.9	529	1	US-08-172-931B-14	Sequence 14, App1
40	210.5	6.9	529	1	US-08-706-037-25	Sequence 25, App1
41	210.5	6.9	529	2	US-09-005-397-25	Sequence 25, App1
42	210.5	6.9	529	2	US-09-032-315-5	Sequence 5, App1
43	210.5	6.9	529	2	US-08-993-318A-5	Sequence 5, App1
44	210.5	6.9	529	4	US-09-399-886-5	Sequence 5, App1
45	210.5	6.9	529	4	US-09-396-260-5	Sequence 5, App1

ALIGNMENTS

RESULT 1
US-09-401-476-4
; Sequence 4, Application US/09401476
; Patent No. 6168936
; GENERAL INFORMATION:
; APPLICANT: Wang, Huming
; TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes
; FILE REFERENCE: GC584
; CURRENT APPLICATION NUMBER: US/09/401,476
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Stachybotrys charatum
US-09-401-476-4

Query Match	100.0%	Score 3067	DB 4	Length 572
Best Local Similarity	100.0%	Pred. No. 2e-289		
Matches 572	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MFKHTLGAALSLFLNSNAVQASVPETSPATGHLFRVAQISPOYPMFTVPLPIPVKQ	60	
DB	1	MFKHTLGAALSLFLNSNAVQASVPETSPATGHLFRVAQISPOYPMFTVPLPIPVKQ	60	
QY	61	PLRLTVTPVNGQELWYVEVEIKPPTHQVYDPLGSAVLVDGMSGPTFOVPRGVETVVR	120	
DB	61	PLRLTVTPVNGQELWYVEVEIKPPTHQVYDPLGSAVLVDGMSGPTFOVPRGVETVVR	120	
QY	121	FINNAEPNVSVHLHGSRSRAFDGMAEDITEPGSFQYYPNROSARTLWYHDAHMTA	180	
DB	121	FINNAEPNVSVHLHGSRSRAFDGMAEDITEPGSFQYYPNROSARTLWYHDAHMTA	180	
QY	181	ENMYRGOAGLYMLTDPEDALNLPSPGYGFDPILTLTSKOYTANGNLVTNGELNSWGD	240	
DB	181	ENMYRGOAGLYMLTDPEDALNLPSPGYGFDPILTLTSKOYTANGNLVTNGELNSWGD	240	
QY	241	VIHVNGQPMFVKVPEPKYFRFLDAAVSFGLYFADTDAIDTRLPEKVIASDSGLLEH	300	
DB	241	VIHVNGQPMFVKVPEPKYFRFLDAAVSFGLYFADTDAIDTRLPEKVIASDSGLLEH	300	
QY	301	PAPTSLLIYMAEYEVFEPDSYAGKTIELRNIGSIGTIDTDYDNDKXWRFVAD	360	
DB	301	PAPTSLLIYMAEYEVFEPDSYAGKTIELRNIGSIGTIDTDYDNDKXWRFVAD	360	
QY	361	DTQOPDTSVVPANLRDVPFSPPTTNTPROFRFGTGTWTINGVAFADVONRLLANVPV	420	
DB	361	DTQOPDTSVVPANLRDVPFSPPTTNTPROFRFGTGTWTINGVAFADVONRLLANVPV	420	

QY 421 TVERMELINANGMTHPHIHLVDFKVISRTSGNNARTVWPYESGLKDVYMLGRRETVV 480
DB 421 TVERMELINANGMTHPHIHLVDFKVISRTSGNNARTVWPYESGLKDVYMLGRRETVV 480
QY 481 EAHYAFPPGYVMHCHNLHEDHDMMAAFNATVLPDYGATVFPVDMELMOARPELG 540
DB 481 EAHYAFPPGYVMHCHNLHEDHDMMAAFNATVLPDYGATVFPVDMELMOARPELG 540
QY 541 EFOAGSGFVSQAVTERIOTMAEYRPPAADE 572
DB 541 EFOAGSGFVSQAVTERIOTMAEYRPPAADE 572

RESULT 2
US-09-401-476-2
; Sequence 2, Application US/09401476
; Patent No. 6168936
; GENERAL INFORMATION:
; APPLICANT: Wang, Huang
; TITLE OF INVENTION: No. 6168936el Phenol Oxidizing Enzymes
; FILE REFERENCE: GC584
; CURRENT APPLICATION NUMBER: US/09/401,476
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Stachybotrys chartarum
US-09-401-476-2

Query Match 65.5%; Score 2010; DB 4; Length 583;
Best Local Similarity 65.4%; Pred. No. 1,1e-186;
Matches 383; Conservative 71; Mismatches 110; Indels 22; Gaps 10;

QY 1 MKRHTGAAALSL-LFNSNAVQASPPV--ETSPATGHLKRVQISQIPMTVPLPIP 57
DB 1 MTSQAIGAVLGLAVIGSSVDARSVAGRSTDMPSG-LTKROTQSLPPLALVEVPLPIP 59
QY 58 VKOPRLTVNPVNGOETIWEVEIKPPTHQVYRDLGASADLVGYDGMSPPTPOVPGVET 117
DB 60 LKAPN-TVPNPTGEDIILYEMELRPSHQIYDLEPANNVGYDGMSPGPTIIVPKTES 118
QY 118 VYRFINNAA--APNSVHLGHSFSAFADGMAEDITEPGSEKDYYPNROSARTLWYHDA 175
DB 119 VYRFVNGENTSPNSVHLGHSFSAFADGMAEDITEPGSEKDYYPNROSARTLWYHDA 178
QY 176 MHTAENAVRGAGLYMLTDPAEDALNLSGIGEDIPMLISKOYANGNLVTNGELN 235
DB 179 MSTAEANAVMGAGVYIMIDDPADALNLSGIGEDIPMLISKOYANGNLVTNGELN 238
QY 236 SFMGADVIHNGOWPFRKNEPRKYRERFLDAVSRSGLYEADTDALDPRKVIASDS 295
DB 239 STMGADVIHNGOWPFRKNEPRKYRERFLDAVSRSGLYEADTDALDPRKVIASDS 298
QY 296 GLEHHPADTSLIYISMAEREVEVDFSDYAGKYTELRNLGSGIGITDIDYDNTDKVMR 355
DB 299 GLEGGVDDIOLYISMAEREVEVDFSDYAGKYTELRNLGSGIGITDIDYDNTDKVMR 357
QY 356 FVYVADDTTPODTSVDPANLADVPP-----SPTTNPTRQFRGRTGTWTINGVAF 407
DB 358 FVYVDEVLESPTSEVPANLADVPPPEGGMNDPANPTDD--ETFTFGRANGQMTINGVTF 415
QY 408 DVONRLLANVPQGVTERMELINANGMTHPHIHLVDFKVISRTSGNNARTVWPYE-SGL 466
DB 416 DVENRLLRNVPRDVTVEIMRLENNNSNGMTHPHIHLVDFKVISRTSG--AGVPEAAGL 472
QY 467 KDVYMLGRRETVVVEAHYAFPPGYVMHCHNLHEDHDMMAAFNATVLPDYGATVFPV 526
DB 473 KDVYMLGRRETVVVEAHYAFPPGYVMHCHNLHEDHDMMAAFNATVLPDYGATVFPV 532

QY 527 PMELMOARPELGEFQAGSGFVSQAVTERIOTMAEYRPPAADE 572
DB 533 PMELMOARPELGEFQAGSGFVSQAVTERIOTMAEYRPPAADE 578

RESULT 3
US-09-032-315-2
; Sequence 2, Application US/09032315
; Patent No. 5985818
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: LACCASE MUTANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5985818 No. 5985818disk of No. 5985818th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,315
; FILING DATE: 27-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 5200,200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-032-315-2

Query Match 7.5%; Score 228.5; DB 2; Length 499;
Best Local Similarity 22.8%; Pred. No. 1,1e-13;
Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

QY 55 IPEVKOPRLTVNPVNGOETIWEVEIKPPTHQVYRDLGASADLVGYDGMSPPTPOVPG 114
DB 2 IGPVAD-LITMAA-----VSPDGRSQAVVNVNGIPGGLINGNG 41
QY 115 VETVYRFINNAA--APNSVHLGHSFSAFADGMAE-----DITEPGSE-KDY 160
DB 42 DREFQVLVIDLNLHTMLKSTSIHMHGFFQGT--NMADGPAFINQCPISGHSFLYDF 99
QY 161 PNRQSARTLWYHDAHHTAENAVRGAGLYMLTDPAEDALNLSGIGEDIPMLISKO 220
DB 100 PDQ--AGTWYHSH-LSTOYCDGLRGRPFVYDNDPAADLYVDN-----DDVT 151
QY 221 YTANGMLVTTNGELNSFWGVDVHNGO-----PWPEKNEP-RKYRERFLDAVSR 270
DB 152 YHVAAL-----GAPFLGADATILNCKGSPSTTTADLSVISTPCKRFRFLVLS 207
QY 271 SE-----GLYFADDAIDTRLPEFKVIASDSGLEHHPADTSLIYISMAEREVEV 323
DB 208 NYFESIDGNMHTIIEEDSINTA-----PLVVDISQIFAQORSFVLE-AN 251
QY 324 YAGKYTELR--NLG--SGISGIGITDIDYDNTDKVMR--VVADDTTPODTSVDPAN 375

Db 252 QAVDNYWIRANPNFGVGTGINS-----ALLRYDGA-AAVEPTTQTSTAPLNEY 303
Qy 376 DVPEPSTPTNTPR-----OFRFGRTGPTWTINGVAFADYONRLANVPVGTV 424
Db 304 NL-HPLVTTAVPGSPVAGVDLAINMAFNPNFTINGASFTPTVPVLLQIISGAONA 362
Qy 425 WELINAGN-----GWTPIHILHVDKVISRTSGNNAATVPYE 463
Db 363 QDLPSSGSVYSLPSNADIEISFPATAAPGAPHPHGHAFVYV-RSAGS---TYNYND 418
Qy 464 SGL-KDYVWLG---RRETVVEAHYAPFGVYMFHCHNLHEDHDMMAFNATVLPDYG 519
Db 419 NPFRDVSSTGTPPAAGDNTVIRFR-TDNPQPMFLHCHIDHLEAGFAVVF-AEDIDP--- 473
Qy 520 NATVFDPMEEELW 532
Db 474 --VASANPVQAM 484

RESULT 4
US-08-993-318A-2
; Sequence 2. Application US/08993318A
; Patent No. 5998353
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Anders
; APPLICANT: Svendsen, Allan
; APPLICANT: Schneider, Palle
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Cherry, Joel
; TITLE OF INVENTION: LACCASE MUTANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59983530 No. 5998353disk of No. 5998353th America
; STREET: 405 Lexington Avenue
; CITY: New York
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,318A
; FILING DATE: December 18, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valera A.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 5032.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELE: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-993-318A-2

Query Match 7.5%; Score 228.5; DB 2; Length 499;
Best Local Similarity 22.8%; Pred. No. 1.1e-13;
Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

Qy 55 IIPVQOPRLTVNPNVNGGIWYVEIKRFTHQVYPDLSADLYDYGSPGPTPQVPRG 114
Db 2 IGPVAD--LTIINAA-----VSPDGSROAVVYVNGGPGGLITGNMG 41
Qy 115 VETVYAFINNAE-----APNSVHLHGSFRAAFDGMAE-----DITEPGSF-KDYVY 160

Db 42 DRFQLVNIDNLTNHTMLKSTSIHMHGFQKGT--NNADCPAFINQCPISGSHFLYDFQV 99
Qy 161 PNRQSAFTLMDHHAHHTENAVRCAGLYMLTDAEDLNLPSGIGEDIDMILLTSKQ 220
Db 100 PEO--AGTEWYHSH-LSTQYCDLGRFVYDPNDPADLYDVDN-----DQVITLVDW 151
Qy 221 YTANGNLVTTNGELNSFWGDVHVHNCQ-----PWPKNVEP-RKYRFRFLDAVSR 270
Db 152 YHVAKL-----GPAFLGADATLNLNGKGRSPSTTADLSVISTPCKRYRFLVSLSDP 207
Qy 271 SF-----GLYPADDAIDTRLPEKVIASDGLLEHPADTSLLYISMAERYVDFDSD 323
Db 208 NYTFSIDGHNMFTIETDSINTA-----PLVVDISIQIFAAQRYSEVLE-AN 251
Qy 324 YAKGTIELR---NLG---GSIGIGITDIDYNTDKVAF---VYADDTQDTSVSPANLR 375
Db 252 QAVDNYWIRANPNFGVGTGINS-----ALLRYDGA-AAVEPTTQTSTAPLNEY 303
Qy 376 DVPEPSTPTNTPR-----OFRFGRTGPTWTINGVAFADYONRLANVPVGTV 424
Db 304 NL-HPLVTTAVPGSPVAGVDLAINMAFNPNFTINGASFTPTVPVLLQIISGAONA 362
Qy 425 WELINAGN-----GWTPIHILHVDKVISRTSGNNAATVPYE 463
Db 363 QDLPSSGSVYSLPSNADIEISFPATAAPGAPHPHGHAFVYV-RSAGS---TYNYND 418
Qy 464 SGL-KDYVWLG---RRETVVEAHYAPFGVYMFHCHNLHEDHDMMAFNATVLPDYG 519
Db 419 NPFRDVSSTGTPPAAGDNTVIRFR-TDNPQPMFLHCHIDHLEAGFAVVF-AEDIDP--- 473
Qy 520 NATVFDPMEEELW 532
Db 474 --VASANPVQAM 484

RESULT 5
US-09-399-886-2
; Sequence 2. Application US/09399886
; Patent No. 6140092
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Anders
; APPLICANT: Svendsen, Allan
; APPLICANT: Schneider, Palle
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Cherry, Joel
; TITLE OF INVENTION: LACCASE MUTANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 61400920 No. 6140092disk of No. 6140092th America
; STREET: 405 Lexington Avenue
; CITY: New York
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/399,886
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,318
; FILING DATE: December 18, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valera A.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 5032.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655

TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 499 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-399-886-2

Query Match 7.5%; Score 228.5; DB 4; Length 499;
 Best Local Similarity 22.8%; Pred. No. 1,1e-13;
 Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

55 IPPVQPRLLTVNPNVNGEIIWEYVEIKPFTHQVYVPLGASADLVGSDGSGPFTQVPRG 114
 DB 2 IGPVAD--LTITNA-----VSPDGFSSRAVYVNGGTPGLITGMNG 41
 QY 115 VETVVRFINNAE-----APNSVHLGSSRAAFDGMAE-----DITEPGSF-KDYXX 160
 DB 42 DRFOLNVIDNLNHTMLKSTSIHMGFFQKGT--NMADGPAPINOCPISSGHSFLYDFQV 99
 QY 161 PNRGASRLMTHDHAMHTAENAYRGQAGLYMLTPADALMLPSGYGDFDIPMLITSQ 220
 DB 100 PDQ--AGTFWYHSH--LSTQYCDGLGPRVYVYDNDPADLYVDN---DQVITLVDW 151
 QY 221 YTANGNLVTTNGELNSFWGDVYIHVNGQ-----PWFKNVEP-RKYRFRFLDAVSR 270
 DB 152 YHVAAKL-----GPAFLGADATLINGKGRSPSTTADLSVITGKRFRFLVSLSCDP 207
 QY 271 SF-----GLYFADTDAIDTRLPRKVIASDSGLLEHPADTSLIYISMAERYEVVDFSD 323
 DB 208 NYTFSIDGHNMHTIETDSINTA-----PLVDSIOIFAQORYSFYLE-AN 251
 QY 324 YAGKTIELR--NLG--GSIGIGTDTDYDNTDKVMRF--VVAADTQPDTSVVPANLR 375
 DB 252 QAVDNTWIRANPNFGVNGFTGINS-----ALIRYDGAALVEPTTQTSTAPLNEV 303
 QY 376 DVFPSPPTTNTPR-----OFRGRTGPTWTINGVAFADYONRLANVPVGTVER 424
 DB 304 NL-HPLVTTAVGSPVAGVDLAIMAFNENGTNFINGASTPTTPVYLLDIIISGAQA 362
 QY 425 WELINAGN-----GWTPIHILHVDKVISRTSGNNAKRYMPE 463
 DB 363 ODLPSSGSVYSLPSNADIEISFPATAAAGAPHPRHLHGHAFAVY--RSAGS---TVYND 418
 QY 464 SGL-KDYVWLG---RRETVVVAHAAPRPGVYMFCHNLIHEDHDMMAFNATVLPDGY 519
 DB 419 NPIFDVYSTGTTPAAGDNTYTRF--TDNPGWFLHCHIDHLEAGFAVYF-AEDIPD--- 473
 QY 520 NATVEVDPMEEML 532
 DB 474 --VASANPVPOAM 484

RESULT 6
 US-09-396-260-2
 Sequence 2, Application US/09396260
 Patent No. 6184015
 GENERAL INFORMATION:
 APPLICANT: Svendsen, Allan
 APPLICANT: Xu, Feng
 TITLE OF INVENTION: LOCASE MUTANTS
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NO. 61840150 No. 6184015disk of No. 6184015th America, Inc.
 STREET: 405 Lexington Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10174
 COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/396,260
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/032,315
 FILING DATE: 27-FEB-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Rozek, Carol
 REGISTRATION NUMBER: 36,993
 REFERENCE/DOCKET NUMBER: 5200, 200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-878-9655
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 499 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-396-260-2

Query Match 7.5%; Score 228.5; DB 4; Length 499;
 Best Local Similarity 22.8%; Pred. No. 1,1e-13;
 Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

55 IPPVQPRLLTVNPNVNGEIIWEYVEIKPFTHQVYVPLGASADLVGSDGSGPFTQVPRG 114
 DB 2 IGPVAD--LTITNA-----VSPDGFSSRAVYVNGGTPGLITGMNG 41
 QY 115 VETVVRFINNAE-----APNSVHLGSSRAAFDGMAE-----DITEPGSF-KDYXX 160
 DB 42 DRFOLNVIDNLNHTMLKSTSIHMGFFQKGT--NMADGPAPINOCPISSGHSFLYDFQV 99
 QY 271 SF-----GLYFADTDAIDTRLPRKVIASDSGLLEHPADTSLIYISMAERYEVVDFSD 323
 DB 208 NYTFSIDGHNMHTIETDSINTA-----PLVDSIOIFAQORYSFYLE-AN 251
 QY 324 YAGKTIELR--NLG--GSIGIGTDTDYDNTDKVMRF--VVAADTQPDTSVVPANLR 375
 DB 252 QAVDNTWIRANPNFGVNGFTGINS-----ALIRYDGAALVEPTTQTSTAPLNEV 303
 QY 376 DVFPSPPTTNTPR-----OFRGRTGPTWTINGVAFADYONRLANVPVGTVER 424
 DB 304 NL-HPLVTTAVGSPVAGVDLAIMAFNENGTNFINGASTPTTPVYLLDIIISGAQA 362
 QY 425 WELINAGN-----GWTPIHILHVDKVISRTSGNNAKRYMPE 463
 DB 363 ODLPSSGSVYSLPSNADIEISFPATAAAGAPHPRHLHGHAFAVY--RSAGS---TVYND 418
 QY 464 SGL-KDYVWLG---RRETVVVAHAAPRPGVYMFCHNLIHEDHDMMAFNATVLPDGY 519
 DB 419 NPIFDVYSTGTTPAAGDNTYTRF--TDNPGWFLHCHIDHLEAGFAVYF-AEDIPD--- 473
 QY 520 NATVEVDPMEEML 532
 DB 474 --VASANPVPOAM 484

[illegible]

```

Db      252 QAVDNYWIRANPNFGNNGVFTTGINS-----AIIIRYCGAAAVBEPTQTQTSTAPLNEV 303
OY      376 DVPEPSPTTNNPR-----QEFRRGTGTTWINGVAEADYONRLLANVPYGTVER 424
Db      304 NL-HPVITTAAPGSPVAGVDLAINMAFNNGTNFPIINGASFTPPVYPLDIIISGAONA 362
OY      425 WELINAGN-----GWTPIIHILVDEKVISRTSGNNAFTVMPYE 463
Db      363 QDLPLSGSVSYSLPSFNADIEISFPATAAPGAPRPHFLHGHAEFAVY-RSAGS---TVYNYD 418
OY      464 SGL-KDQVWLG---RRTYVVEAHYAFAFCVYVYFHCHNLHEDHDMMAFNATVLPDGY 519
Db      419 NPFRDVAVSTETPAAGNVITRFR-TDNPQPMFLCHIDFHLEAGRAVVF-AEDIPD--- 473
OY      520 NATFVDPMEELW 532
Db      474 --VASANPVRQAW 484

RESULT      8
US-08-462-484-2
Sequence 2, Application US/08462484
Patent No. 5667531
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie Sue
APPLICANT: Xu, Feng
APPLICANT: Dalborge, Henrik
APPLICANT: Schneider, Palte
APPLICANT: Aaslyng, Dorrit A.
TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 56675310 No. 5667531disk of No. 56675311th America, Inc
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,484
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/441,147
FILING DATE: 15-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4185.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 878 9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Polyporus pinsitus
US-08-462-484-2

```

Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

55 IPVKQRLVTNPVNGQELIWEVEIKPETHQVYPLDLSADLVGDGMSPGPTFOVPRG 114
 23 IGVAD--LTITNA-----VSPDGFQRQAVVNGGTPGLITGNMG 62
 115 VETVRFINNAE-----APNSVHLHGSFRAFDGMAE-----DTEPGSF-KDYXX 160
 63 DRQLVNIJNLHTHTWKSSTIMHGFQKGT--NMADGAFINOCPISSGSHFLVDFOY 120
 161 PNRQSARTLWHDHMHITAEANVYRQAGLYMLTDPADLNLPSGGEEDIMILLTSKO 220
 121 PDG--AGTEWYHSH--LSTQYCDGLRGPEVYVDNDPADLDYVDN-----DQVITLVDM 172
 221 YTAGNMLVTTNGELNSFWGVDIVHNGQ-----PWPKNPEP-RKYRFRFLDAVSR 270
 173 YHVAKL-----GPAFLGADATLNLNGKGRSPSTTADLSVISTPGKRYRFLVLSLSDP 228
 271 SF-----GLYFADTDATDRLPFKVIASDGLLEHPADTSLIYISMAERYEVEFSD 323
 229 NYFFSIDGHNMFTIETDSINTA-----PLVDSIQIFAAQRYEFLVLE-AN 272
 324 YAGKTIELR---NLG--GSIGIGTDTDYDNTDKVME--VVAADTQDPTSVFANLR 375
 273 QAVDNYWIRANPFGVNGFTGINS-----ALLRYDGAALVEPTTQTSTAPLNEY 324
 376 DVFPSPPTNTPR-----QFRFGTGPWTNGVAFADVONRLLANPVGTVER 424
 325 NL-HPLVTTAVPGSPVAGVDLAINNAFNGNFTNGSTFPPTVPLQIISGAQNA 383
 425 WELINAGN-----GWTPIHILHVDKFIYSRTSGNNARTWMPYE 463
 384 QDLPLSGSVYSLPSNADIEISFPATAAGCAPHPHILGHAFVAV--RSAGS---TYNYND 439
 464 SGL-KDVMWLG---RRETVVEAHVAPFGVYMFHCHNLHEDHDMMAFNATVLDYGY 519
 440 NPJFRDVSSTGTPAACDNTVIRER-TDNQGPWFLHCHIDHLEAGRAVAV-AADIDP--- 494
 520 NATVFDPMELM 532
 495 --VASANPVQAW 505

RESULT 9
 US-08-441-147-2
 ; Sequence 2, Application US/08441147
 ; Patent No. 5770418
 ; GENERAL INFORMATION:
 ; APPLICANT: Yaver, Debbie Sue
 ; APPLICANT: Xu, Feng
 ; APPLICANT: Dalboge, Henrik
 ; APPLICANT: Schneider, Palle
 ; APPLICANT: Aaslyng, Dorrit A.
 ; TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 57704180 No. 5770418disk of No. 5770418th America, Inc.
 ; STREET: 405 Lexington Avenue, Suite 6400
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/441,147
 ; FILING DATE: 15-May-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:

NAME: Lowney, Karen A.
 REGISTRATION NUMBER: 31,274
 REFERENCE/DOCKET NUMBER: 4185.010-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 867 0123
 TELEFAX: 212 878 9655
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 520 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORGANISM SOURCE:
 ORGANISM: Polyporus pinusitus
 US-08-441-147-2

Query Match 7.5%; Score 228.5; DB 1; Length 520;
 Best Local Similarity 22.8%; Pred. No. 1,2e-13;
 Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

55 IPVKQRLVTNPVNGQELIWEVEIKPETHQVYPLDLSADLVGDGMSPGPTFOVPRG 114
 23 IGVAD--LTITNA-----VSPDGFQRQAVVNGGTPGLITGNMG 62
 115 VETVRFINNAE-----APNSVHLHGSFRAFDGMAE-----DTEPGSF-KDYXX 160
 63 DRQLVNIJNLHTHTWKSSTIMHGFQKGT--NMADGAFINOCPISSGSHFLVDFOY 120
 161 PNRQSARTLWHDHMHITAEANVYRQAGLYMLTDPADLNLPSGGEEDIMILLTSKO 220
 121 PDG--AGTEWYHSH--LSTQYCDGLRGPEVYVDNDPADLDYVDN-----DQVITLVDM 172
 221 YTAGNMLVTTNGELNSFWGVDIVHNGQ-----PWPKNPEP-RKYRFRFLDAVSR 270
 173 YHVAKL-----GPAFLGADATLNLNGKGRSPSTTADLSVISTPGKRYRFLVLSLSDP 228
 271 SF-----GLYFADTDATDRLPFKVIASDGLLEHPADTSLIYISMAERYEVEFSD 323
 229 NYFFSIDGHNMFTIETDSINTA-----PLVDSIQIFAAQRYEFLVLE-AN 272
 324 YAGKTIELR---NLG--GSIGIGTDTDYDNTDKVME--VVAADTQDPTSVFANLR 375
 273 QAVDNYWIRANPFGVNGFTGINS-----ALLRYDGAALVEPTTQTSTAPLNEY 324
 376 DVFPSPPTNTPR-----QFRFGTGPWTNGVAFADVONRLLANPVGTVER 424
 325 NL-HPLVTTAVPGSPVAGVDLAINNAFNGNFTNGSTFPPTVPLQIISGAQNA 383
 425 WELINAGN-----GWTPIHILHVDKFIYSRTSGNNARTWMPYE 463
 384 QDLPLSGSVYSLPSNADIEISFPATAAGCAPHPHILGHAFVAV--RSAGS---TYNYND 439
 464 SGL-KDVMWLG---RRETVVEAHVAPFGVYMFHCHNLHEDHDMMAFNATVLDYGY 519
 440 NPJFRDVSSTGTPAACDNTVIRER-TDNQGPWFLHCHIDHLEAGRAVAV-AADIDP--- 494
 520 NATVFDPMELM 532
 495 --VASANPVQAW 505

RESULT 10
 PCR-US95-07536-2
 ; Sequence 2, Application PC/TUS9507536
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; APPLICANT:
 ; TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07536
FILING DATE: 15-June-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/265,534
FILING DATE: 24-June-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4185.204-MO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 878 9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Polyporus pinitus
PCT-US95-07536-2

Query Match 7.5%; Score 228.5; DB 5; Length 520;
Best Local Similarity 22.8%; Pred. No. 1.2e-13;
Matches 126; Conservative 70; Mismatches 212; Indels 145; Gaps 29;

55 IPVKQPRLVTVNPGVGEIWEVEIKPPTHQVYPLDGSADLVGYDGMSPGFQVPRG 114
23 IGVVAD-LITVNA-----VSPDGFROAVVAVGTPGGLITGCMG 62
115 VETVRFINNAE-----APNSVHLHGSFRAAPDGMAE-----DITPGRSF-KDYVY 160
63 DRQLANVLDLHTHTWKSISIMHGFQKGT--NMADGAFINQCRISGSHSLYFOV 120
161 PNRQSARTLWYHDHAMHTAENYRGAGLYMLTDPADALNPSGYGEFDIPMLTSKQ 220
121 PDG--AGTEWYHSH-LSTQYCDGLRGPFVYVDPNDPAADLYVDN-----DDFVITLVW 172
221 YTANGNLVYTTNGELNSWGVYIHNGQ-----PWPKNVEP-KRYRFLDAVSR 270
173 YHVAKL---GAFFPLGATLNLNGGRSPSTTTADLSVYIPGKRYRFLVLSLSCDP 228
271 SF-----GLYFADDAIDTRLPEKVIASDGLLEHPADTSLIYIMARREYVDFSD 323
229 NYTFESIDGHMTIETDSITA-----PLVYDSIQIFAAQRYSVLE-AN 272
324 YAKTIELR--NLG--GSIGIGITDTDYDNTDKVMRF--VYVADTTOPDTSVVPANLR 375
273 QAVDNDWIRANPNFNGVNGFTGINS-----ALLRYDGAAVEPTTGTSTAPLNEV 324
376 DVFPPEPTTTPR-----QRFEGRTGPTWTNGVAFADVONRLLANVPGTVER 424
325 NL-HPLVTTAVPSPVAGVDLAINNAFNNGNFEFNGTSFTPTVPLQIISGONA 383
425 WELINAGN-----GWTPIHITHLWDFKVIYSTSGNMATVAPYE 463
384 QDLPSGSVYLSNMDIETSPATAAPGAPRPHLGHAFVAV-RSAS--TYVNYD 489
464 SGI-KDYVWLG---RRETVVYEAHYAFPGVYMFCHNLIHEDHMAAENATVLPDGY 519
440 NPJFRDVSSTGTGAAGDNVTIRFR-TDNPCPWLHCHIDHLEAGFAIVE-AEDIDP--- 494

520 NATVVDPMELW 532
495 --VASANPVPQAW 505

RESULT 11
US-08-940-661A-2
Sequence 2, Application US/08940661A
Patent No. 5795760
GENERAL INFORMATION:
APPLICANT: BERKA, Randy Michael
APPLICANT: BROWN, Stephen H.
APPLICANT: XU, Feng
APPLICANT: SCHNEIDER, Pallo
APPLICANT: OXENB LL, Karen M.
TITLE OF INVENTION: PURIFIED MYCELIOPHTHORA LACCASES AND NUCLEIC
ACIDS OF INVENTION: ACIDS ENCODING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 57957600 No. 5795760disk of No. 5795760th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940,661A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/441,146
FILING DATE: 15-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 4184.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 620 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-940-661A-2

Query Match 7.4%; Score 226; DB 1; Length 620;
Best Local Similarity 22.8%; Pred. No. 2.8e-13;
Matches 155; Conservative 80; Mismatches 255; Indels 190; Gaps 37;

3 KHTLGAALSLLEFNSNVOASPYETSPATGHLFKRVAQISPOYPMETVPL---PIPPVK 59
2 KSFISAATLLVGLITPVAAP-----PST-----PEQRDLVPTTETEEAAVK 45
60 QPRLVTNPNVNGEIMY-----YEVE-----IKPPTHQVYPLDGSAD-LVGYDGM 104
46 AROOSCTPSN-RACMTWDGVDINTDYEVSDPDGVVAPYT---LTLEVDNMTGPDGV 100
105 P-----GPTFVPPRGVETVAFINNAE-PNSVHLHGSFRAA--PQGMADIT 150
101 KEKVMVNNISITGPTIPADMGDTIQVTVINNLLETNGTSIMHGHGKTNLHDG-ANGIT 159
151 E---PGSFQDYVYPNRQSARTLWYHDHAMHTAENYRGAGLYMLTDPADALNPSG 206

```

Db 160 ECEIPKGRKRYRFRKAQGYGTSWYHS---HFSAQYG-NGVGAIOINGPA--SLPYDTD 213
QY 207 YGFEIDPMLITSKOYANGNLVTTNGELNSFGCDVIHVNGQ-----PWPEKRV---EP 256
Db 214 LGVFPI-----SDYYSSADELVELTKNSGAFESDNYLFNCTAHNPETGEGEYANVLTTPG 269
QY 257 RKYRFRFLDAAVSRSGLYPAD---TDAIDTRLPEFVYIASDSGLLEHPADTSLYISMAE 313
Db 270 RRRRLRLINTSVENHFOVSLVNHNTTIIAADVYPVANMAYDS-----LFLGVGQ 318
QY 314 RYEVVEFSDYAK-----TIELRLNGSIGIGTD-----TDYD 348
Db 319 RYDVVLEASTPGNWFNVTFFGGGLCGSRRNDYPAIFHYAGAPGPPDEKAPVDHN 378
QY 349 NTD-KVMRFVADDTQOPDTSVYPAULRDYPPSPPTTNPQRFRTGTPTWING-----403
Db 379 CIDLPLMKPVYADVPLSGAKRPDNTLDVTL--DTTGPTLF-----VKNVNSAIN 428
QY 404 -----VAFADYQNRLLANVPVG--TVERWELINANGMT-----HPIH 439
Db 429 IDMGRRVVDYVLTON---TSFPPGYNIVE-----VNGADQMSYWLIENTDPCAPFTLPHPMH 481
QY 440 IHLVDEKVIISR-----TSGNNARTVM--PYESGL-----KDVWLGRRRETVVEAHYA 485
Db 482 LHHGHDYVLGRSPDESPPASNERHVPDPARDAGLLSGANPYRVDYTMLPAGWVVL-AFRA 540
QY 486 PFFGVYMFCHNLIHEDHDMMAAFNATVLPDYGYNATVFVDPMEL-----WQARPYE 538
Db 541 DNGGAMLFCHIAMHVSGLGVYLLERADDLRGAVSDADADDLRLCADMRRRYWPTNPYP 600
QY 539 -----LGEFOAGSGQFSVQA 553
Db 601 KDSGLKHRWVEEGEVLVKA 620

RESULT 12
US-09-083-485-2
; Sequence 2, Application US/09083485
; Patent No. 5948121
; GENERAL INFORMATION:
; APPLICANT: DORRIT Aaslyng
; TITLE OF INVENTION: Laccases with Improved Dyeing
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59481210 No. 5948121disk of No. 5948121th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,485
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valeta
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 4639,204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 620 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```

US-09-083-485-2

Query Match          7.4%; Score 226; DB 2; Length 620;
Best Local Similarity 22.8%; Pred. No.2.8e-13;
Matches 155; Conservative 80; Mismatches 255; Indels 190; Gaps 37;

QY 3 KHTLGAALSLLENSNAVQASPVETSPATGHLFKHVAQISPOLYPMFTVPL---PIPPYK 59
Db 2 KSFTISAATLLVGLTFSVAAP-----PST-----PQRRDLVITREEAAYK 45
QY 60 QPRLVTNPVNGOEIY-----YEVE-----IKPETHQYVPLDGSAD-LVYIDGS 104
Db 46 AROQSCNTPSN-RACWTIDGYDINTDYEDVSPDYGVRPRYH---LTLTEDMNTGPDGVV 100
QY 105 P-----GPTQVPRGVEYVREFINNAEA-DNSYVHLGHSFRAA--PFGMAEDIT 150
Db 101 KEKVYMLVNNSTIOPITFADMGDTIOYTVINNLETNGTSHIMHGLHOKTLYHDC-ANGIT 159
QY 151 E-----PGEFKDYVYPPNROSARTLMDHDMHITAEANAVRQAGLYMLTDPADALNLPSC 206
Db 160 ECEIPKGRKRYRFRKAQGYGTSWYHS---HFSAQYG-NGVGAIOINGPA--SLPYDTD 213
QY 207 YGFEIDPMLITSKOYANGNLVTTNGELNSFGCDVIHVNGQ-----PWPEKRV---EP 256
Db 214 LGVFPI-----SDYYSSADELVELTKNSGAFESDNYLFNCTAHNPETGEGEYANVLTTPG 269
QY 257 RKYRFRFLDAAVSRSGLYPAD---TDAIDTRLPEFVYIASDSGLLEHPADTSLYISMAE 313
Db 270 RRRRLRLINTSVENHFOVSLVNHNTTIIAADVYPVANMAYDS-----LFLGVGQ 318
QY 314 RYEVVEFSDYAK-----TIELRLNGSIGIGTD-----TDYD 348
Db 319 RYDVVLEASTPGNWFNVTFFGGGLCGSRRNDYPAIFHYAGAPGPPDEKAPVDHN 378
QY 349 NTD-KVMRFVADDTQOPDTSVYPAULRDYPPSPPTTNPQRFRTGTPTWING-----403
Db 379 CIDLPLMKPVYADVPLSGAKRPDNTLDVTL--DTTGPTLF-----VKNVNSAIN 428
QY 404 -----VAFADYQNRLLANVPVG--TVERWELINANGMT-----HPIH 439
Db 429 IDMGRRVVDYVLTON---TSFPPGYNIVE-----VNGADQMSYWLIENTDPCAPFTLPHPMH 481
QY 440 IHLVDEKVIISR-----TSGNNARTVM--PYESGL-----KDVWLGRRRETVVEAHYA 485
Db 482 LHHGHDYVLGRSPDESPPASNERHVPDPARDAGLLSGANPYRVDYTMLPAGWVVL-AFRA 540
QY 486 PFFGVYMFCHNLIHEDHDMMAAFNATVLPDYGYNATVFVDPMEL-----WQARPYE 538
Db 541 DNGGAMLFCHIAMHVSGLGVYLLERADDLRGAVSDADADDLRLCADMRRRYWPTNPYP 600
QY 539 -----LGEFOAGSGQFSVQA 553
Db 601 KDSGLKHRWVEEGEVLVKA 620

RESULT 13
US-08-939-218A-2
; Sequence 2, Application US/08939218A
; Patent No. 5981243
; GENERAL INFORMATION:
; APPLICANT: BERKA, Randy Michael
; APPLICANT: BROWN, Stephen H.
; APPLICANT: XU, Feng
; APPLICANT: SCHNEIDER, Palle
; APPLICANT: OXENB LL, Karen M.
; APPLICANT: AASLYNG, DORRIT A.
; TITLE OF INVENTION: PURIFIED MYCELIOPTHORA LACCASES AND NUCLEIC
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59812430 No. 5981243disk of No. 5981243th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor

```

CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/939,218A
FILING DATE: 29-SEPT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol E.
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 4184.120-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 620 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-939-218A-2

Query Match 7.4%; Score 226; DB 2; Length 620;

Best Local Similarity 22.8%; Pred. No. 2,8e-13; Matches 155; Conservative 80; Mismatches 255; Indels 190; Gaps 37;

QY 3 KHTLGAALSLFNSNAVQSPVSPATGHLEKRAVQISPOQPMFTVTL---PIRPVK 59
DB 2 KSTISATLVLGLTSPVAAP---PST-----PEGRDLVPTTEREAAVK 45
QY 60 QPRILVTNPNVNGEIMW-----YEVE-----IKPETHOYVPDLGSAD-LVYDGMS 104
DB 46 ARQSCPTSPN-RACWTDGVDIMTDEVDSDPDGCVARPT---LTLVEDNMTGPRGVV 100
QY 105 P-----GPTFOVPRGVETVAFFINNAE-PNSVHLHGSFSRAA--FDGMAEDIT 150
DB 101 KEKVMVLNNSIIGPTLFADMGDTIQVIVINNLETNGTSIMHGHGKGTMLHGG-ANGIT 159
QY 151 E-----GSGFDYVYYPNROSARTLMYHDMHMTAENAYRQAGLYMLTDPAEDALNPSG 206
DB 160 ECPTRPGGKRVKFAQOYGTSTYHS--HFSQYQ-NGVGAIQINGPA--SLPYDITD 213
QY 207 YGEFDIPMILTSQYTAGNMLVTTNGELNSFGVDIVHNGQ-----PWPCKV---EP 256
DB 214 LGVFP-----SDYYSADDELVELTKNSGAPFSNIVLFNGTAKRPGESEYANVTITPG 269
QY 257 RKYRFRFLDAVSRSEGLYFAD--TDAIDTRLFVKVIASDSGLLEHPADTSLIYSMAE 313
DB 270 RRHRLRLINTSVENHEQVSLVNHMTIADAMVFNAMTVDS-----LFLGVGQ 318
QY 314 RYEVVPEFSDYAK-----TIELNLGSGTIGIGD-----TDYD 348
DB 319 KYDVYLEASTPGMYWENVTFFGGLLCGSRNYPRAIFHYAGAPGPPDEGKAPVDHN 378
QY 349 NFD-KVMRFVADDTQPDTSVVPANLADVPPSPNTNTPROGFRGTGPTWTNG---- 403
DB 379 CLDLPLNKKPVVADVPLSGAKRDNLDVTL-DTGTGLF-----VMKYGSAIN 428
QY 404 -----VAFADVONRLLANVPVG--TVERMELINAGNGT-----HPIH 439
DB 429 IDMGKPVVDVLTQN--TSFPPGYNIVE-----VNGADQMSYMLIENDGAPFTLPPMH 481
QY 440 IHLVDFKVISR-----TSGNNAARTVM--PYESGL-----KQVVMGRETIVVEAHYA 485
DB 482 LHGHDFVLGRSPDESPASNERHYVDPARDAGLLSGANPYRVDVTLPAFGWVVL-AFRA 540

QY 486 PFGVYMFHCHNLIEDHDMMAFNATVLPDYGYNATVFPMEEL-----WQARPYE 538
DB 541 DNPGLMFHCHIAHWVSGGLGVYLERADDLKGAVSDADDLRLCADMRRTWPTNPYP 600
QY 539 -----LGEFOAQSGQSFVQA 553
DB 601 KSDGLKHRWVEEGEWLVKA 620

RESULT 14

US-08-462-484-8
Sequence 8, Application US/08462484
Patent No. 5667531

GENERAL INFORMATION:

APPLICANT: Yaver, Debbie Sue
APPLICANT: Xu, Feng
APPLICANT: Dalboge, Henrik
APPLICANT: Schneider, Paule
APPLICANT: Aaslyng, Dorrit A.
TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 56675310 No. 5667531disk of No. 5667531th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,484
CLASSIFICATION: 435
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/441,147
FILING DATE: 15-MAY-1995

ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4185.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 878 9655
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 511 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-484-8

Query Match 7.3%; Score 223.5; DB 1; Length 511;
Best Local Similarity 21.7%; Pred. No. 3,5e-13;
Matches 117; Conservative 60; Mismatches 227; Indels 135; Gaps 21;

QY 34 HLEKRVQISPOYMPFVPLPIRPVKOPRLTVTPVNGEIMWYEVEIKPETHOYVDPDLG 93
DB 15 HSEGRVS-----AALGPVTD-LTISN-----GIVSPDGF 42
QY 94 SADLVYDGSPGPTFOVPRGVETVAFFINNAE-----APNSVHLHGSFSRAAFDGMAE 147
DB 43 TRAAVLANGVFPGLITGNKGDNFQINVIDNLNETMLKSTSIHMHGFQKGT--NMADG 100
QY 148 -----DIEPQSGFKRIYYPNROSARTLMYHDMHMTAENAYRQAGLYMLTDPAEDA 200
DB 101 AAEVNOCPATIGNSFL-YDFTATDQACTFWYHSH-LSTOYCDGLRGPMYVDPSPHADL 158

RESULT 15
 US-08-441-147-8
 : Sequence 8, Application US/08441147
 : Patent No. 5770418
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Yaver, Debbie Sue
 :
 : APPLICANT: Xu, Feng
 :
 : APPLICANT: Dalboge, Henrik
 :
 : APPLICANT: Schneider, Palte
 :
 : APPLICANT: Aaslynd, Dorrit A.
 :
 : TITLE OF INVENTION: PURIFIED POLYPORUS LACCASES AND
 :
 : TITLE OF INVENTION: NUCLEIC ACIDS ENCODING SAME
 :
 : NUMBER OF SEQUENCES: 10
 :
 : CORRESPONDENCE ADDRESS:
 :
 : ADDRESSEE: NO. 5770418b NO. 5770418disk OF NO. 5770418th America, Inc.
 : STREET: 405 Lexington Avenue, Suite 6400

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Version #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,147
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lowmeyer, Karen A
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4185,010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 878 9655
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 511 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
OS-08-441-147-8

```

Query Match	7.3%;	Score 223.5;	DB 1;	Length 511;
Best Local Similarity	21.7%;	Pred. No. 3.5e-13;		
Matches 117;	Conservative 60;	Mismatches 227;	Indels 135;	Gaps 21.

[illegible]

Search completed: October 3, 2002, 10:54:53
Job time: 37 sec

THIS PAGE BLANK (USPTO)